Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Anthony S. Ransdell

Entitled:

Investigating the Biosynthetic Pathway of Polyacetylenic Natural Products in *Fistulina hepatica* and *Echinacea purpurea*

For the degree of Master of Science

Is approved by the final examining committee:

Robert E. Minto
Chair

Eric C. Long

Lei Li

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Robert E. Minto

Approved by: Martin J. O'Donnell 10/31/2012
Head of the Graduate Program Date

Graduate School Form 20
(Revised 9/10)

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:
Investigating the Biosynthetic Pathway of Polyacetylenic Natural Products in *Fistulina hepatica* and *Echinacea purpurea*

For the degree of Master of Science

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Anthony S. Ransdell
Printed Name and Signature of Candidate

10/31/2012
Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

INVESTIGATING THE BIOSYNTHETIC PATHWAYS TO POLYACETYLENIC NATURAL PRODUCTS IN *FISTULINA HEPATICA* AND *ECHINACEA PURPUREA*

A Thesis

Submitted to the Faculty

of

Purdue University

by

Anthony S. Ransdell

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2012

Purdue University

Indianapolis, Indiana

I would like to dedicate this work to my wife, Kassie; my parents, Dan and Annie Ransdell; and my brothers, Aaron and Austin Ransdell. Without your continued support, understanding, and unconditional love this would not be possible. It is because of you that I am who I am today.

ACKNOWLEDGEMENTS

I owe a great debt of gratitude to many people who have helped me along the way. I would especially like to thank my mentor Dr. Robert Minto for his wisdom, patience, and support. I would also like to thank my committee members for their willingness to help me through this journey and their valued insight both inside and outside of the classroom. I would like to thank all of the members of the Blacklock research group, particularly Selene Hernandez-Buquer for her help and friendship. I would also like to thank Michael Shepard Jr. for his patience and assistance in both natural product isolation and NMR data acquisition and analysis.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

LIST OF ABBREVIATIONS

aa	amino acid
ACP	acyl carrier protein
Amp	ampicillin
cDNA	complimentary deoxyribonucleic acid
CM	complete minimal media
CoA	coenzyme A
COSY	correlation spectroscopy
DCM	dichloromethane
dex	dextrose
DMF	dimethylformamide
DMOX	dimethyloxazoline
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotides
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
EST	expressed sequence tag
FA	fatty acid
FAME	fatty acid methyl ester
FAPy	fatty acid pyrrolidide
gal	galactose
GC-MS	gas chromatography-mass spectroscopy
gDNA	genomic deoxyribonucleic acid

His	histidine
HSQC	heteronuclear single quantum correlation
KIE	kinetic isotope effect
LB	Luria-Bertani broth
Leu	leucine
mRNA	messenger ribonucleic acid
MS	Murashige and Skoog
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide
NMR	nuclear magnetic resonance
nt	nucleotide
PC	phosphatidylcholine
PCR	polymerase chain reaction
PPM	plant preservative mixture
PUFA	polyunsaturated fatty acid
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
RT-PCR	reverse transcription-PCR
SDM	site-directed mutagenesis
SOC	super optimal broth
TAG	triacylglycerol
TLC	thin-layer chromatography
UV	ultraviolet

ABSTRACT

Ransdell, Anthony S. M.S., Purdue University, December 2012. Investigating the Biosynthetic Pathways to Polyacetylenic Natural Products in *Fistulina hepatica* and *Echinacea purpurea*. Major Professor: Robert E. Minto.

Polyacetylenic natural products, compounds containing multiple carbon-carbon triple bonds, have been found in a large collection of organisms. Radiochemical tracer studies have indicated that these bioactive metabolites are synthesized from fatty acid precursors through a series of uncharacterized desaturation and acetylenation steps. To date, there are three main pathways believed to be involved in acetylenic natural product biosynthesis. However, it is apparent that the crepenynic acid pathway is the origin of a vast majority of the known plant and fungal acetylenic products.

This investigation provides concrete evidence that the polyacetylenic natural products found in the fungus *Fistulina hepatica* and the medicinal plant species *Echinacea purpurea* are biosynthesized from crepenynic acid. Through heterologous expression in *Yarrowia lipolytica*, two acetylenases capable of producing crepenynic acid were identified from *E. purpurea*. Furthermore, heterologous expression of two diverged desaturases isolated from *F. hepatica*, uncovered a Δ12-acetylenase and the first multifunctional enzyme capable of Δ14-/Δ16- desaturation and Δ14-acetylenation.

CHAPTER 1. INTRODUCTION

Fatty acid desaturases are a family of enzymes capable of carrying out oxidative dehydrogenation chemistry on established acyl chains. Over the years, a series of diverged desaturases has emerged that have the ability to insert a variety of functional groups into fatty acid acyl chains. One particular example of these diverged desaturases is the alkyne-forming subclass known as acetylenases. These enzymes further desaturate pre-existing double bonds in an acyl chain to produce the carbon-carbon triple-bonded product.

Over the last several decades, polyacetylenic natural products from plants, fungi, and aquatic organisms were discovered to possess a range of bioactivities. These bioactivities include, but are not limited to, anti-tumor, HIV-inhibitory, and immunosuppressive properties. The unique biological activities possessed by polyacetylenic compounds have made them of great interest to the medical and pharmaceutical industries. However, there is still much to be discovered about the dehydrogenation pathways responsible for the biogenesis of these natural products.

This thesis describes investigations of polyacetylenic natural product biosynthesis in *Fistulina hepatica* and *Echinacea purpurea*. Provided that many acetylenic metabolites originate from fatty acids and fatty acid precursors, the introduction begins with a brief overview of fatty acids. The fatty acid desaturase family is described, highlighting the different classes of desaturases and their characteristics, as the formation of the acetylenic bond is thought to undergo dehydrogenation in an analogous fashion to that of integral membrane desaturases. Lastly, the introduction concludes with a focus on acetylenase enzymes and provides the basis for why understanding the biosynthesis of acetylenic natural products is important.

1.1 Overview of Fatty Acids

1.1.1 What are Fatty Acids?

Fatty acids (FAs) are a cornerstone of existence since they play crucial roles in the biological structures and functions that support life. Fatty acids are essential biological components that make up the building blocks of all acyl lipids. FAs and their complex lipid derivatives act as: 1) energy storage reserves; 2) major components of membrane bilayers for both the cell and individual organelles; 3) precursors of biologically active compounds, such as hormones and secondary messengers; and 4) protein modifiers capable of imparting proper folding and anchoring of membrane proteins through acylation [1, 2]

A completely unmodified FA consists of two parts: 1) a hydrocarbon chain with a methyl group at one end and 2) a carboxylic acid functional group at the other end (Figure 1). *Saturated* fatty acids contain acyl chains with only methylene groups, where the hydrocarbon portion of the acyl chain is saturated with hydrogens. Fatty acid modification pathways can introduce additional functional groups into the acyl chain, however. These chemical modifications include a large number of functional groups, for example fluoro, hydroxyl, double, and triple bonds to name a few. When functional groups such as epoxides, alkenes, and alkynes are present, these fatty acids are considered *unsaturated* because their hydrocarbon chains no longer contain the maximum complement of hydrogens.

In plants, *de novo* FA biosynthesis largely occurs either in the chloroplasts or the stroma of plastids, whereas in most other eukaryotes biosynthesis occurs in the cytosol [3]. The process of fatty acid biosynthesis is carried out by either a multisubunit (prokaryotic) or a multifunctional (eukaryotic) fatty acid synthase complex that uses the precursors acetyl-CoA and malonyl-CoA to generate saturated fatty acids with chain lengths of 16-18 carbons. However, modifications of *de novo* fatty acid synthesis

products, through elongation and β-oxidation processes, can provide acyl chains ranging from 10-30 carbons in length.

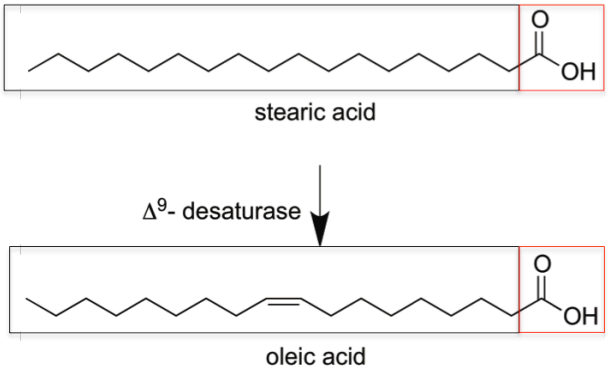


Figure 1: Structure of fatty acids. Steric acid is a saturated fatty acid whereas oleic acid is unsaturated. The installation of the double bond is carried out by a desaturase enzyme. The boxes in red and black outline the carboxylic acid and hydrocarbon portions of the fatty acid, respectively.

1.1.2 Fatty Acid Naming and Nomenclature

Three types of names are regularly used to identify FAs: 1) common names; 2) IUPAC nomenclature; and 3) lipid shorthand notation. The most abundant monounsaturated fatty acid found in nature is systematically named *cis*-9-octadecenoic acid. However, the same fatty acid is frequently referred to by its common name oleic acid or its lipid shorthand notation $18{:}1^{9}$. In the shorthand notation $x{:}y{:}^{z}$, 1) the number of carbon atoms in the acyl chain are given by x; 2) the number of double bonds present are enumerated by y; and 3) the position(s) of modifications in the acyl chain and the geometric configuration of the double bond, when necessary, are designated by the superscript 'z'. By using the (Δ) shorthand notation, oleic acid can be abbreviated as $18{:}1^{\Delta 9c}$. When using the (Δ) notation for more exotic fatty acids, such as (*Z*)-octadec-9-en-12-ynoic acid (Figure 2), the notation is the same with one exception. Although the acetylenic functional group contains two units of unsaturation and is annotated in superscript form, it is not counted as a double bond. Consequently, the short hand notation for (*Z*)-octadec-9-en-12-ynoic acid is $18{:}1^{\Delta 9c,12a}$. Other shorthand notations will not be discussed as only the (Δ) notation is used in this thesis.

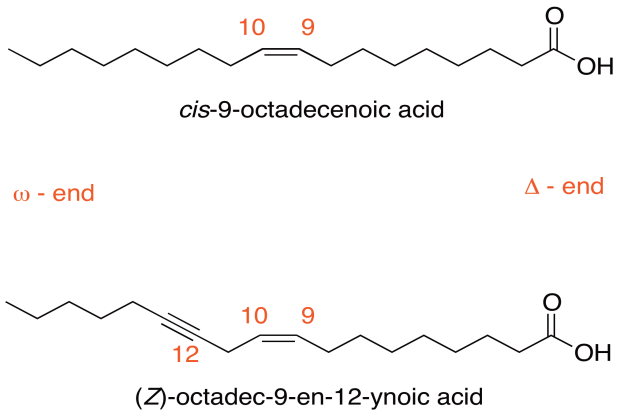


Figure 2: Shorthand notation for fatty acids. In this notation, the Δ represents the position of unsaturation relative to the carboxyl end (Δ- end) of the acyl chain. In the case of the shorthand notation for (*Z*)-octadec-9-en-12-ynoic acid (18:1$^{\Delta 9c,12a}$) the *cis* double bond and the triple bond are located at positions C9 and C12 relative to the carboxyl end of the acyl chain, respectively. An alternate naming system counts from the methyl or ω-end.

1.2 Fatty Acid Desaturases

Fatty acid desaturases are a family of enzymes capable of catalyzing the dehydrogenation of pre-existing fatty acyl chains. *Sensu stricto* this family acts to remove adjacent hydrogen atoms from consecutive methylene groups in order to generate a double bond in the acyl chain. Although the double bonds occur almost exclusively in the *cis* configuration, *trans* isomers do exist in Nature as a minor proportion of unsaturated species. The chemistry of double-bond formation requires two electrons, two protons, and molecular oxygen. The electrons serve as reducing equivalents while activated O_2 helps facilitate the cleavage of the highly stable C-H bond and acts as the final electron acceptor, ultimately becoming reduced to water during the process.

This family of enzymes is found in both prokaryotic and eukaryotic organisms, and is known to be involved in the regulation of membrane fluidity. Furthermore, desaturases are responsible for the formation of polyunsaturated fatty acids (PUFAs), known precursors to a large number of biologically active molecules in plants and animals. For example the eicosanoid, arachidonic acid (20:4$^{\Delta 5c,8c,11c,14c}$), is a highly unsaturated fatty acid that can be used to generate hormones like prostaglandins,

thromboxanes, and leukotrienes, as well as other biologically-active compounds, such as pheromones and growth regulators [1, 4].

A survey of this enzymatic family has shown that desaturases carry out a broad range of dehydrogenation chemistries using various substrates and producing a multitude of unsaturated products. Desaturases are capable of modifying four distinct classes of acyl substrates: acyl-CoAs, acyl-ACPs, complex acyl-lipids, and sphingolipids [5]. In addition to substrate specificity, these desaturases have also been found to have different regioselectivities. For example, front-end desaturases only catalyze dehydrogenation of C_4-C_8, while other desaturases express regioselectivities, primarily Δ9-, Δ12-, and Δ15.

Many attempts have been made at categorizing the desaturase family into further subclasses. Categorization has been done on the basis of substrate specificity, regioselectivity, and cell location. However, a more specific approach is to categorize desaturases based on their amino acid sequences. In using this approach, two evolutionarily distinct groups of desaturase enzymes have emerged.

1.3 The Two Classes of Desaturases

The smaller of the two classes includes the soluble desaturases. To date, soluble desaturases have only been identified in plants, and are exclusively located in the plastid, where they have been shown to act solely on acyl substrates esterified to an acyl carrier protein (ACP). Crystallographic and enzymological studies have been made possible by the soluble nature and relative ease of isolation of this group of desaturases. In 1993, Fox *et al.* used optical and Mössbauer spectroscopies to study a stearoyl-ACP Δ9-desaturase from castor (*Ricinus communis)*. Based on the spectroscopic data and sequence homology, this study defined the presence of a di-iron cluster and a highly conserved amino acid motif thought to be involved in catalysis [6]. Through crystal

structures of the castor and ivy ACP-desaturases, the participation of two conserved D/EXXH motifs in binding the catalytically essential non-heme di-iron complex was confirmed [7-9].

The second and much more widespread class of desaturases is comprised of integral membrane proteins, which includes enzymes involved in the biosynthesis of exotic fatty acids (see section on diverged FAD2s below) and biologically essential PUFAs. Although the two conserved D/EXXH motifs found in soluble desaturases are absent, three highly conserved histidine motifs (or boxes) have been identified and are thought to be involved in catalysis. Similar to the soluble class, these His boxes are believed to participate in a non-heme di-iron complex. These three histidine motifs contain eight His residues that, upon individual mutation, resulted in the loss of desaturase activity [10]. These eight His residues form three highly conserved boxes of the sequences: $H(X)_{3-4}H$, $H(X)_{2-3}HH$, and $H/Q(X)_{2-3}HH$. Additionally, membrane-bound desaturases contain 4-6 conserved hydrophobic regions, which are proposed to include four α-helical transmembrane domains that act to anchor the enzyme to the membrane bilayer and orient the His boxes on the cytosolic side of the membrane [9, 11, 12] (Figure 3). Further purification and characterization studies of this class of enzymes have been significantly impeded by the inherent instability of these redox-active proteins outside of their native membranous environments.

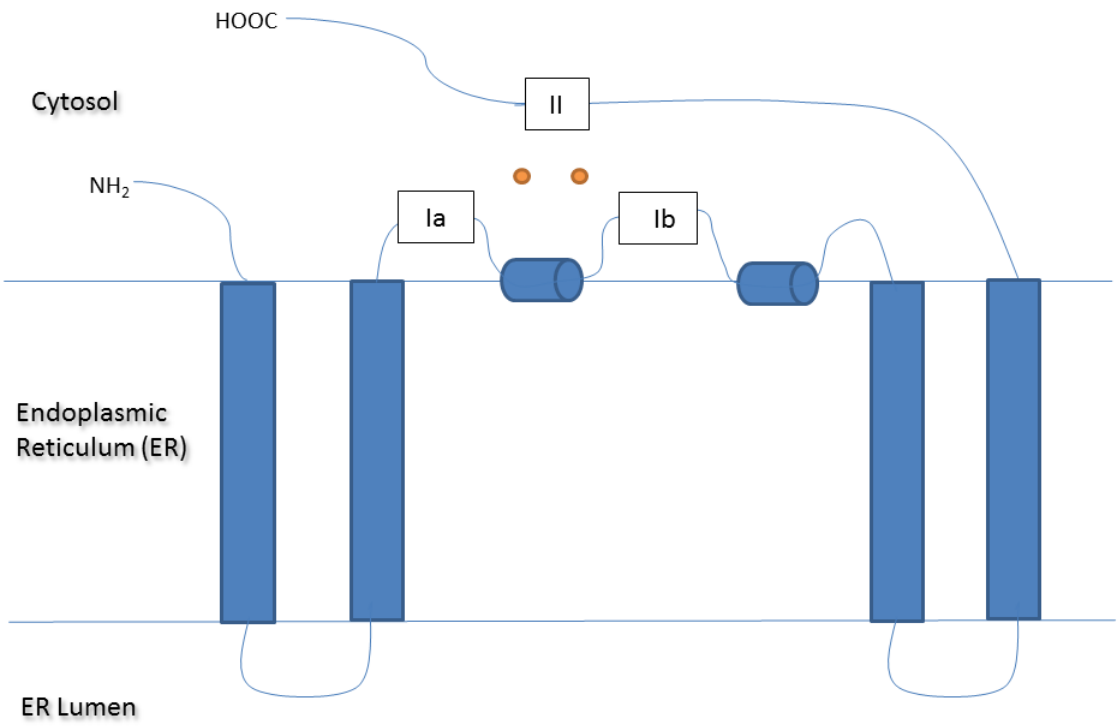


Figure 3: A proposed topological model for a FAD2 enzyme. The membrane-bound enzyme is anchored to the ER by four transmembrane domains. Also illustrated are the highly conserved His boxes that are thought to bind a non-heme di-iron cluster (shown in orange) on the cytosolic side of the bilayer.

A series of investigations on the phylogenetically unrelated *Pseudomonas oleovorans* alkane ω-hydroxylase (AlkB) may help provide insight into the relationship between soluble and membrane-bound desaturation chemistry. AlkB is a bacterial integral protein that requires iron and molecular oxygen for the conversion of octane to n-octanol [13]. Mössbauer spectroscopic studies demonstrated the presence of a non-heme di-iron cluster similar to that found in the soluble stearoyl-ACP Δ9-desaturase from castor, but with spectroscopic parameters consistent with multiple His Ligands [14]. Additionally it was discovered that the di-iron cluster was reduced upon addition of sodium dithionite, but was reoxidized by the addition of O_2 and substrate proving the di-iron cluster's potential for involvement in catalysis [14].

Despite the general lack of AlkB sequence homology with known membrane-bound desaturases, it has been found that AlkB contains the same three His motifs,

which are also positionally conserved relative to the hydrophobic regions [15]. Studies carried out by Shanklin demonstrated a complete loss of hydroxylase activity when each of the eight conserved His residues was mutated to alanine [15]. Taking the Mössbauer and mutagenesis data together it becomes clear that AlkB contains a catalytically active non-heme di-iron cluster similar to that found in the soluble class of desaturases. This more robust enzyme may help bridge the gap between soluble and membrane-bound desaturase catalysis, and serve as an accessible mechanistic and structural model for membrane-bound desaturation.

In order to catalyze dehydrogenation chemistry, all desaturases require an electron source, a flavoprotein, and a conjugating group, which is frequently a carrier protein. The electron transport components required for catalysis are linked to the cellular localization rather than the solubilization of the group [8]. The reducing electrons for plastidal-based, soluble desaturation are provided by NADPH. Reducing equivalents are then transferred to the flavoprotein ferredoxin-$NADP^+$ oxidoreductase and passed down the electron transport chain to the soluble desaturase through the 2Fe-2S ferredoxin electron carrier. However, if the soluble desaturase is located in photosynthetic tissue, the reducing equivalents eminate from photosystem I and pass directly to ferredoxin thus negating the need for NADPH and the flavoprotein ferredoxin-$NADP^+$ oxidoreductase [16].

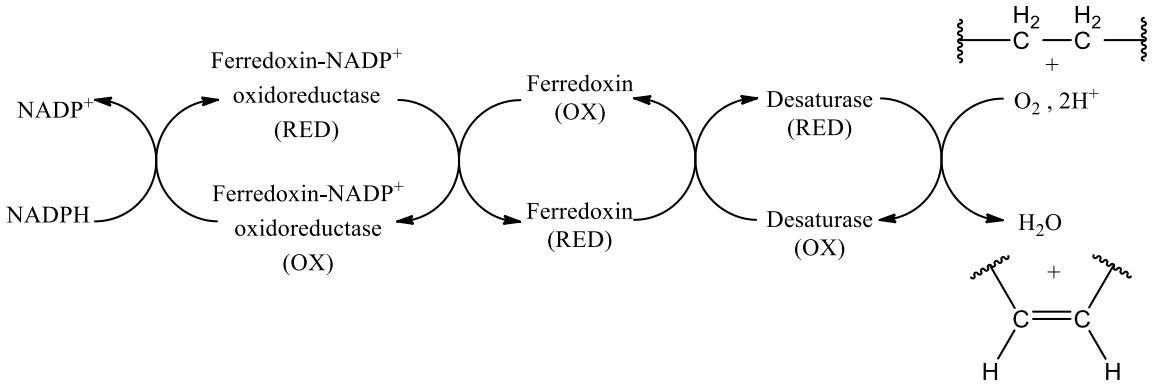


Figure 4: Electron transport system for soluble desaturases. This figure shows the flow of electrons in the plastid from NADPH to the desaturase enzyme.

Conversely, in the case of microsomal systems, the reducing equivalents come from NADH, which are shuttled through the transport system by the flavoprotein cytochrome b_5 reductase. These electrons are passed through the cytochrome b_5 carrier protein to the desaturase, from where they are ultimately transferred to the final acceptor O_2 (Figure 5).

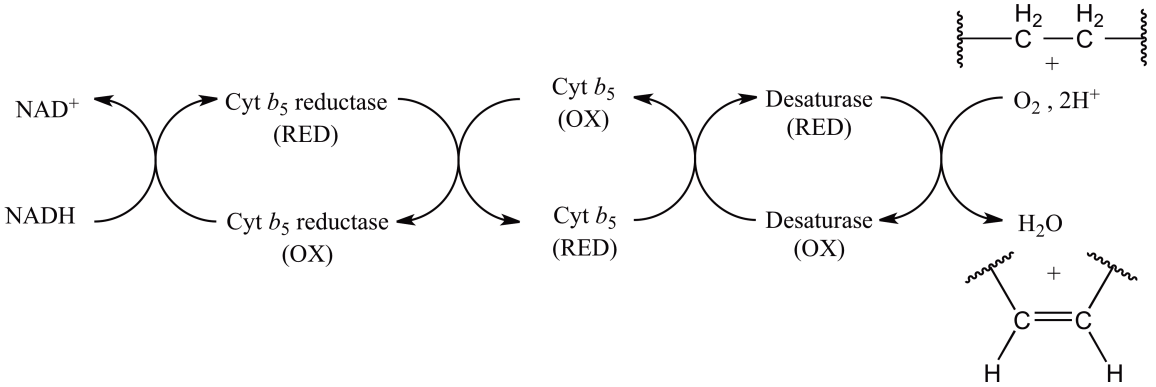


Figure 5: Electron transport system for microsomal desaturation. This figure shows the flow of electrons in the endoplasmic reticulum from NADH to the desaturase enzyme.

1.4 FAD2s

In plants and fungi, 18:2 fatty acids are generated by the membrane-bound desaturase FAD2. The archetypical *fatty acid desaturase-2* (*fad2*) gene encodes a Δ12-desaturase located within the endoplasmic reticulum membrane that desaturates $18{:}1^{\Delta 9c}$ esterified to phosphatidylcholine (PC) to form the product $18{:}2^{\Delta 9c,12c}$. FAD2 generally modifies substrates of 16-18 carbons in length and produces double bonds in the *cis* configuration. The FAD2 family has been shown to contain the characteristic transmembrane domains and highly conserved His boxes that are indicative of the membrane-bound desaturases.

Kinetic isotope effect (KIE) studies have been conducted to try to elucidate the mechanism of desaturation carried out by FAD2 in converting $18{:}1^{\Delta 9c}$ to $18{:}1^{\Delta 9c,12c}$. Heterologous expression of *A. thaliana* FAD2 in *S. cerevisiae* supplemented with [12,12-$^2H_2$]-oleate and [13,13-$^2H_2$]-oleate demonstrated a large primary deuterium isotope effect (k_H/k_D = 7.3 $\pm$ 0.4) corresponding to the C12-H bond cleavage [17]. Furthermore,

it was discovered that the C13-H bond-breaking step was insensitive to deuterium substitution (k_H/k_D = 1.05 $\pm$0.04). A second investigation into the mechanism of FAD2 desaturation was carried out by Behrouzian [18]. The observed KIE of the *Chlorella vulgaris* FAD2, where the values for C12-H and C13-H were k_H/k_D = 7.4 $\pm$ 0.3 and k_H/k_D = 1.08 $\pm$ 0.01, respectively, was nearly identical to the data obtained for the *A. thaliana* homolog. By studying the KIE of stereospecifically labeled oleates, Stymne was able to determine that the pro-(*R*) hydrogen at C12 is the proton abstracted first in the formation of linoleate [19]. Taken together, these data provide evidence for a stepwise mechanism of desaturation, in which the rate-limiting hydrogen abstraction occurs at the pro-(*R*) hydrogen on C12. This is followed by a rapid collapse of the highly unstable intermediate and subsequent loss of a hydrogen attached to the C13 atom.

1.5 Diverged FAD2s

Through broad efforts to identify and investigate FAD2s from plants, fungi, and other organisms, it has been discovered that a range of FAD2 subtypes exists. Many of these enzymes are known as divergent FAD2s, FAD2-like desaturases, or FAD2 variants, for their proficiency in carrying out a wide array of fatty acid modifications while possessing a similar primary sequence to the prototypical enzymes. Diverged FAD2 desaturases have the ability to carry out hydroxylation, epoxidation, conjugation, and acetylenation chemistry with unsaturated fatty acids (Figure 6). These exotic fatty acids are important commercial products. Hydroxy fatty acids can be used as lubricants, in plastics, and in cosmetics, [20, 21] whereas conjugated products are used as drying agents, such as in paints and ink [22].

In 1995, the first FAD2 variant, a Δ12-hydroxylase from castor (*Ricinus communis)*, was described. The hydroxylase was realized to have 67% sequence similarity with the well-studied FAD2 of *Arabidopsis* [23]. When the gene was expressed in transgenic tobacco, an accumulation of low levels of ricinoleic acid ($18{:}1^{\Delta 9c,12OH}$) was observed at 0.1% of the total seed fatty acids. A second cDNA encoding a

Δ12-hydroxylase has since been isolated from *Lesquerella fendleri* [24]. This cDNA showed 71% and 81% identity with the castor hydroxylase and *Arabidopsis* FAD2, respectively. Upon expression of the *L. fendleri* hydroxylase in *Arabidopsis*, the transgenic seeds were observed to accumulate 18:1$^{\Delta 9c,12OH}$, 18:1$^{\Delta 9c,12OH,15c}$, 20:1$^{\Delta 11c,14OH}$ and trace amounts of 20:2$^{\Delta 11c,14OH,17c}$.

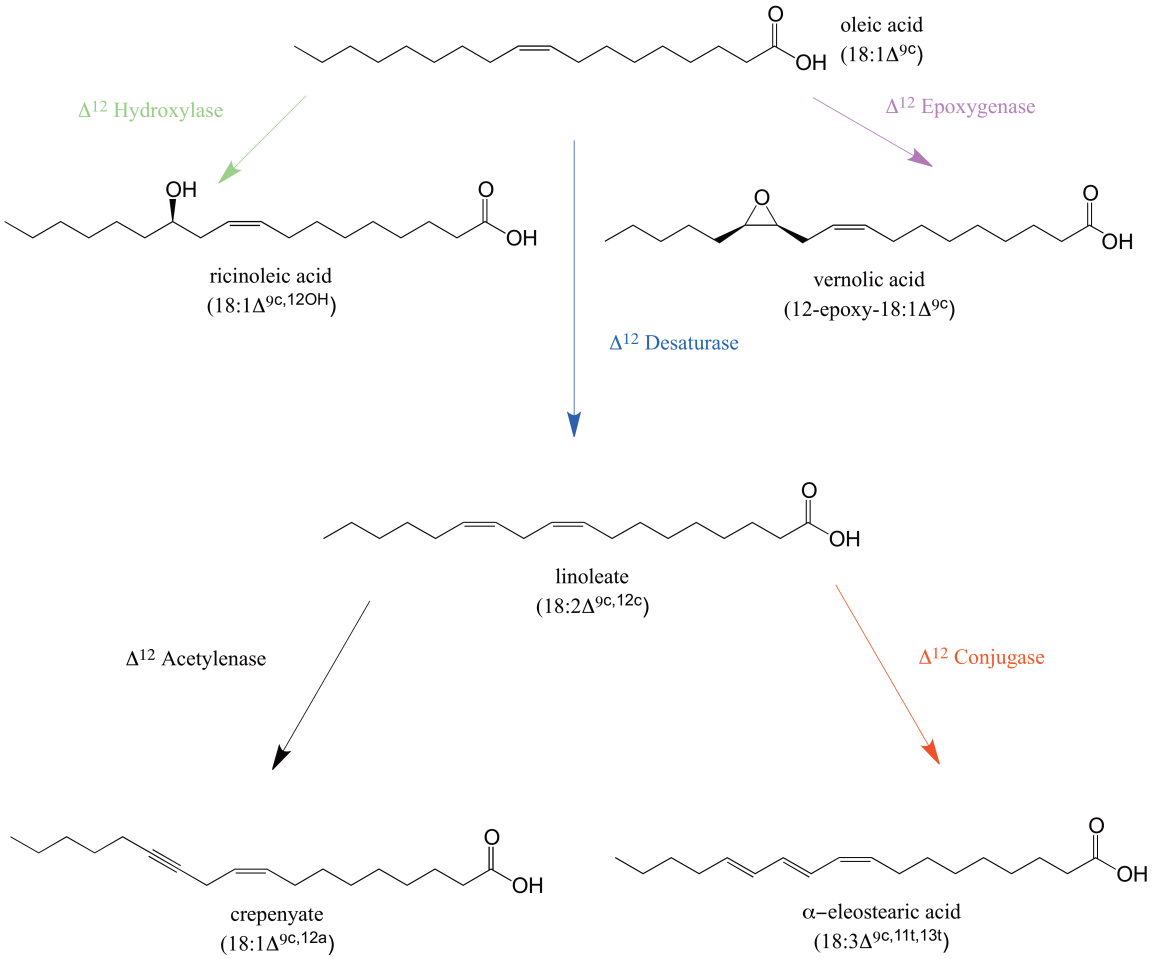


Figure 6: FAD2 and FAD2 variant chemical modifications. The prototypical FAD2 carries out the Δ12 desaturation step indicated by the blue arrow. Meanwhile, FAD2 variants carry out manipulations that result in hydroxyl, epoxy, conjugated, and acetylenic fatty acids represented by the green, pink, red, and black arrows, respectively.

Research into the *Crepis* genus has established that *C. alpina* and *C. palaestina* seed oils contain roughly 70% crepenynic acid (18:1$^{\Delta 9c,12a}$) and 60% vernolic acid (12-epoxy-18:1$^{\Delta 9c}$), respectively [25]. In 1998, the Stymne lab characterized the enzymes responsible for the production of both crepenynic and vernolic acids in these two

species [25]. A full-length cDNA designated *Cpal2* was obtained from *C. palaestina* seeds and the deduced amino acid sequence was found to have 58% identity with the *Arabidopsis* FAD2 and 53% identity with the aforementioned castor Δ12-hydroxylase. *Cpal2* was transformed into *A. thaliana* and the total fatty acids from transgenic seeds accrued up to 15% of 12-epoxy-18:1$^{\Delta 9c}$, indicating that *Cpal2* from *C. palaestina* is a Δ12-epoxygenase [25].

Similarly, a full-length cDNA (*Crep1*) was obtained from *C. alpina* seeds that coded for a protein with 56% sequence homology to *Arabidopsis thaliana* FAD2, 59% with the castor Δ12-hydroxylase, and 81% identity with Cpal2. The fatty acid profile of transgenic yeast expressing the *Crep1* cDNA revealed a 0.3% accumulation of crepenynic acid (18:1$^{\Delta 9c,12a}$). Upon expression of *Crep1* in *Arabidopsis*, the total fatty acids of transgenic seeds contained as much as 25% 18:1$^{\Delta 9c,12a}$ [25]. This landmark study was the first time that a Δ12-acetylenase or Δ12-epoxygenase had been cloned and functionally characterized.

Another example of divergent FAD2 activity can be viewed in the 2002 investigation carried out by Dyer [22]. In this study, it was discovered that two cDNAs, designated *FAD2* and *FADX*, isolated from tung seeds *(Aleurites fordii* Hemsl) were largely homologous to the *Arabidopsis* FAD2. The cDNA indicated as *FAD2* possessed 78% homology and carried out the quintessential FAD2 Δ12 desaturation. However, the *FADX* cDNA was identified as a bifunctional, divergent FAD2 with 69% identity to *A. thaliana* FAD2. The *FADX* gene displayed Δ12-conjugase activity, acting to convert the Δ12 bond of 18:2$^{\Delta 9c,12c}$ into the conjugated product 18:3$^{\Delta 9c,11t,13t}$ [22].

Many of the FAD2 variants contain multifunctional activities. Beyond having Δ12 conjugation activity, FADX performed Δ12 desaturation stereoselectively in the *trans* configuration generating a multitude of desaturation products, ranging from 16:2$^{\Delta 9c,12t}$ to 18:2$^{\Delta 9c,12t}$ to 18:4$^{\Delta 6c,9c,11t,13t}$ [22]. Further investigation into *Crep1* and *Cpal2* exhibited that *Crep1* has both Δ12 *cis* and *trans* desaturation activity, as well as the acetylenase activity described earlier, while the Δ12-epoxygenase was uncovered to also possess a weak *trans* Δ12 desaturation activity [19].

1.6 Acetylenases

A number of acetylenase genes have been isolated from various organisms and functionally characterized. To date, acetylenases universally belong to the membrane-bound desaturase group. They harbor the same four hydrophobic regions characteristic of other membrane desaturases and the three His motifs that are believed to be involved in catalysis.

With many of the same structural features as membrane-bound desaturases, it is believed that the mechanism used for the formation of the alkyne is analogous to the mechanism used in generating an alkene functional group (Figure 7). Evidence supporting this hypothesis was obtained through kinetic isotope studies. Deuterium KIE experiments, conducted through heterologous expression of *CREP1* in *S. cerevisiae*, uncovered a stepwise mechanism for the dehydrogenation of linoleate to crepenynate [26]. The C-H bond at C12 was very sensitive to labeled substitution ($k_H/k_D = 14.6 \pm 3.0$) but, in contrast, the C13 cleavage was relatively unaffected ($k_H/k_D = 1.25 \pm 0.08$). This data provides evidence that the acetylenation reaction proceeds via initial breakage of the C12-H bond, and is followed by a rapid cleavage of the C13-H bond analogous to the Δ12-desaturation mechanism described earlier.

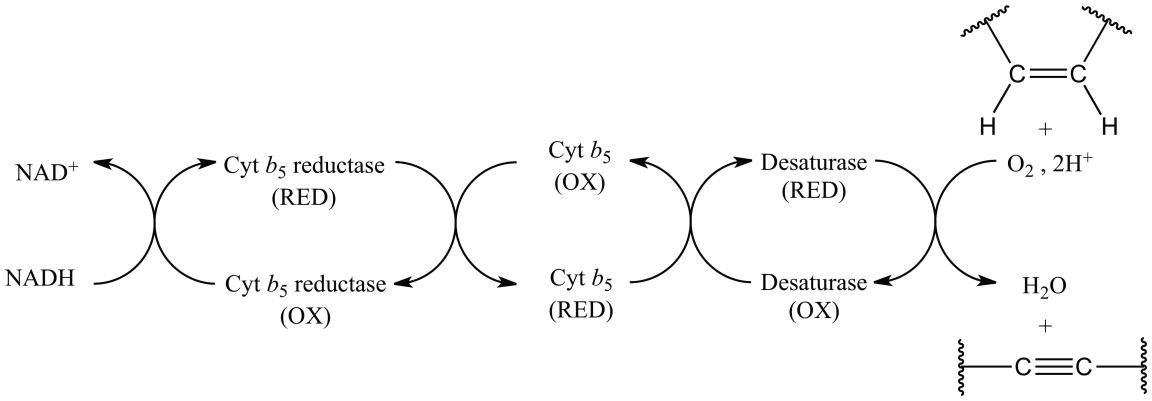


Figure 7: The proposed electron transport system and mechanism for acetylenation.

The isolation and functional characterization of several other plant acetylenases has been made possible by their substantial sequence homology to FAD2 and other known acetylenases. The “typical” FAD2 from parsley has 74% sequence identity with

Arabidopsis FAD2, whereas the translation of the variant *ELI12* gene has 61% [27]. The *ELI12* gene, whose expression is induced by fungal elicitation, demonstrated no activity when expressed in *S. cerevisiae*. In contrast, expression of the gene in soybean displayed the accumulation of both $18{:}1^{\Delta 9c,12a}$ and (14*Z*)-dehydrocrepenynic acid ($18{:}2^{\Delta 9c,12a,14c}$) to a level of 1.4 and 2.2% of the total FAMEs, respectively [28]. This data allowed for the *ELI12* gene to be redesignated *PcACET*.

The functionality of other divergent FAD2s from English ivy (*Hedera helix*), sunflower (*Helianthus annuus*), and *Calendula officinalis* were also investigated [28]. The corresponding FAD2 variant from ivy was found to contain 90% identity with PcACET, and upon expression in soybean embryos the accumulation of $18{:}1^{\Delta 9c,12a}$ and $18{:}2^{\Delta 9c,12a,14c}$ was over 6% of the total fatty acids, combined. The FAD2-related EST from sunflower has only 57% homology to HhACET, but was shown to have approximately 75% amino acid identity to the *Crepis* acetylenase and epoxygenase. Again, the transgenic soybean embryos expressing the diverged desaturase from sunflower was verified to accumulate crepenynate and dehydrocrepenynate. Originally, Fritsche *et al.* functionally characterized the divergent FAD2 from *C. officinalis* as a Δ8, Δ11-desaturase, acting on the $18{:}2^{\Delta 9c,12c}$ substrate and forming the conjugated fatty acid product, calendic acid ($18{:}2^{\Delta 8t,10t,12c}$). However, the characterization was based on a flawed interpretation of UV data of fatty acids isolated from recombinant yeast expression cultures [29]. As the *Calendula* enzyme possessed 91% amino acid sequence identity to the *H. annuus* Δ12-acetylenase, it was reinvestigated and functionally determined to be an acetylenase upon expression in somatic soybean embryos [28].

In addition to plant Δ12-acetylenases, there are examples of Δ12-acetylenases from other eukaryotes, as well as, other acetylenases with different regioselectivities. A fungal Δ12-acetylenase from *Cantharellus formosus*, a moss bifunctional Δ6-acetylenase, and a multifunctional Δ11-acetylenase from the moth *Thaumetopoea pityocampa* have been reported [30-32]. These examples illuminated the potential for fatty acyl acetylenases with different regiospecificities, as is the case with other fatty acid

desaturases. If so, this ever-expanding variation in substrate specificity could contribute to the large diversity of acetylenic natural products.

1.7 Acetylenic Natural Products

Acetylenic natural products are metabolites that contain a carbon-carbon triple bond, and have been unearthed in a variety of organisms ranging from terrestrial plants, mammals, amphibians, and marine algae. As the most frequently encountered acetylenic natural products are polyacetylenes, early research into these compounds was limited by the inherent instability and seeming ease of *in vivo* degradation of acetylenes. Starting in the 1950's and carrying through the 1990's, studies into the synthesis, isolation, and biogenesis of acetylenic natural products were pioneered by the labs of Sir E.H.R. Jones, F. Bohlmann, and J. D. Bu'Lock [33, 34].

The first real insight into the biosynthesis of acetylenic secondary metabolites came when [1-^{14}C]acetate feedings in the fungus, *Polyporus anthracophilus,* revealed the labeling of the odd-numbered carbons in matricaria ester (Figure 8) [33, 35]. In addition, labeled acetylenic natural products have been observed in feeding experiments with [2-^{14}C]-malonate [33, 34]. Given these results, it was proposed that naturally occurring acetylenes are mainly derived from fatty acids, although many details of simple fatty acid synthesis was unknown at the time [33].

The formation of acetylenic natural products is thought to proceed by one of three major biosynthetic pathways: the crepenynic acid pathway, the stearolic acid pathway, and the tariric acid pathway. These pathways differ from one another based on the regioselectivities of desaturation and/or acetylenation of the precursor oleic acid, leading to the formation of crepenynic acid (18:1$^{\Delta 9c,12a}$), stearolic acid (18:0$^{\Delta 9a}$), and tariric acid (18:0$^{\Delta 6a}$), respectively. A survey of the literature appears to suggest that the vast majority of acetylenic metabolites are synthesized via the crepenynic acid pathway (Figure 9). It is believed that the wide range of chain lengths present in acetylenic

metabolites (ranging from C_8-C_{18}) are primarily derived from C_{18} fatty acids that have their chain length altered through oxidative processes as they are processed into the secondary metabolite.

Over the years, feeding experiments with labeled fatty acid precursors, as well as, labeled fatty acids have led to a model of acetylenic metabolite production beginning with basic fatty acids [33-35]. One of the most notable biochemical advancements in the field came in 1998 when the Stymne lab cloned and characterized the first Δ12-acetylenase, Crep1 [25]. Stymne showed that Crep1 uses linoleate as a substrate to form the branchpoint metabolite crepenynic acid by further desaturating the *cis* double bond at the Δ12 position, establishing the relationship between desaturases and acetylenases.

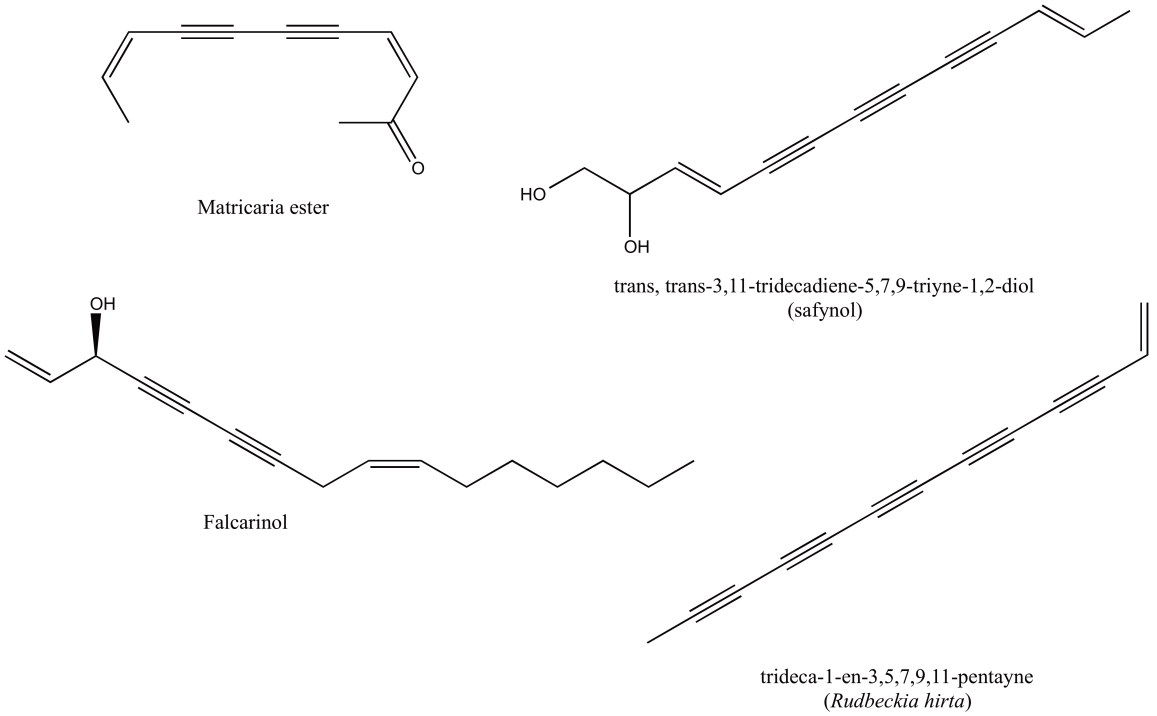


Figure 8: Biologically active acetylenic natural products.

On several occasions, these naturally occurring metabolites have been found to contain biological activities including: anti-fungal, anti-bacterial, anti-viral, and anti-inflammatory properties. An example of these bioactive properties is illustrated by the

finding that the *Carthamus tinctorius* metabolite safynol inhibited the growth of the fungus *Phytophthora drechsleri* [36]. Additionally, three classes of polyacetylenes from *Rudbeckia hirta* have been shown to have photoinduced insecticidal properties toward mosquito larvae and the tobacco hornworm [37]. Furthermore, the acetylenic fatty acid dehydrocrepenynate is an inhibitor of bacterial conjugation, which plays a key role in the spread of bacterial antibiotic resistance [38]. It has also been determined that at lower, dietary concentrations (–)-falcarinol, a polyacetylene isolated from carrots and ginseng, has chemoprotective effects, as verified in a murine colon cancer model [39].

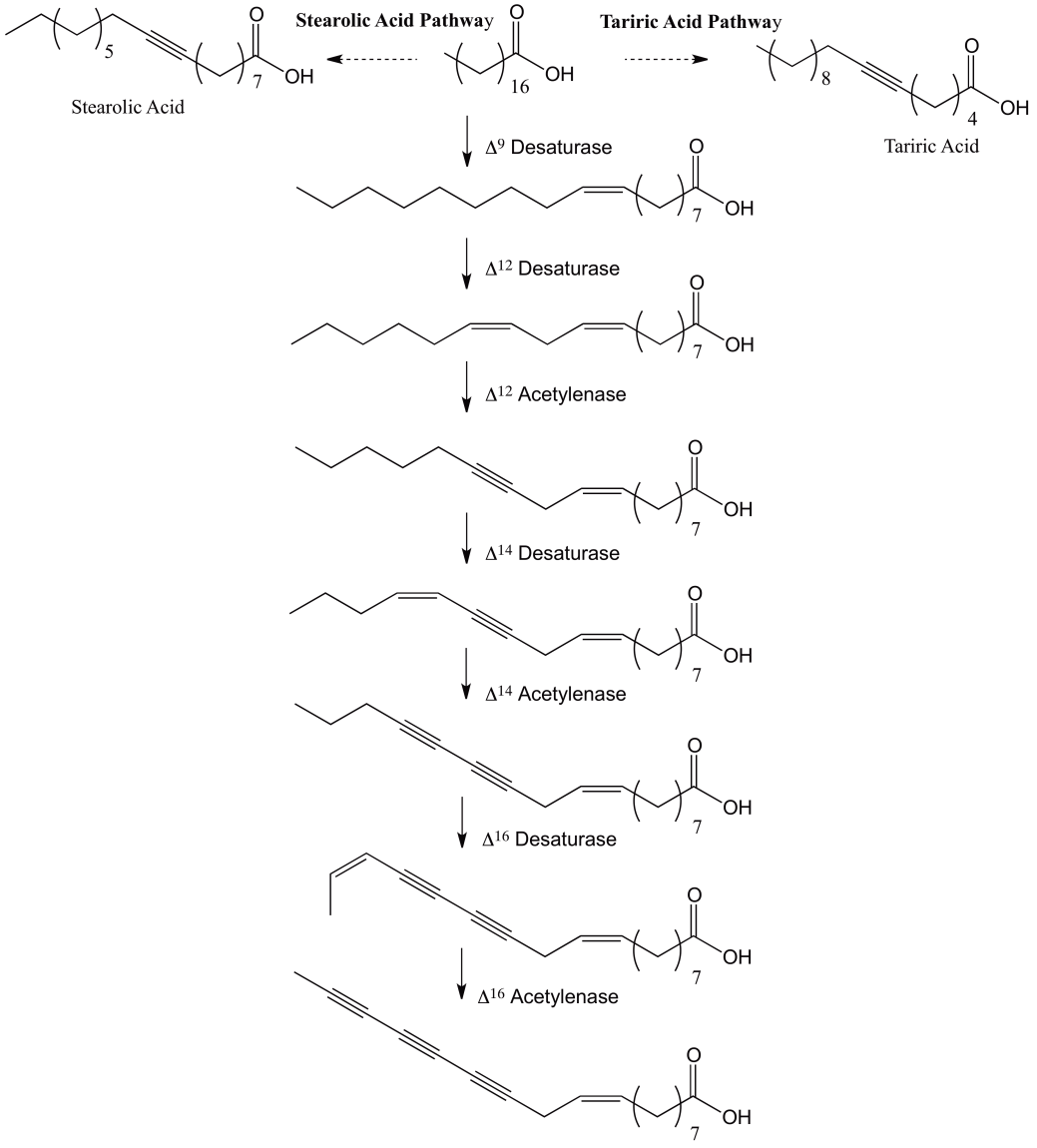


Figure 9: The crepenynic acid pathway as originally proposed by Bu'Lock [40].

Many other examples of polyacetylene bioactivity have been identified over the past couple of decades [34, 41]. However, specific cases are beyond the scope of the current work and therefore, interested readers are directed to the above literature for a more detailed background of these interesting properties. Nevertheless, the biological activities of this class of compounds indicate the significance for research in the area of acetylenic natural products, while there is still much investigation needed to elucidate whether a general mechanism is involved in the biosynthesis of these molecules.

This thesis investigates the dehydrogenation pathways involved in the formation of polyacetylenes found in the fungus *Fistulina hepatica* and the plant *Echinacea purpurea.* Heterologous expression in yeast of gene targets isolated from both *F. hepatica* and *E. purpurea* have led to the production of crepenynic acid, verifying their ability to execute the first step in the formation of polyacetylenic natural products. Additionally, a second gene isolated from *F. hepatica* encoded the ability to perform multiple dehydrogenation steps beyond the formation of crepenynic acid, that are consistent with the hypothetical crepenynate pathway proposed by Bu'Lock. These two studies have provided concrete evidence supporting the hypothesis that the crepenynic acid pathway is involved in the biogenesis of natural acetylenes accumulated by *F. hepatica* and *E. purpurea.*

CHAPTER 2. *FISTULINA HEPATICA*

2.1 Background

Basidiomycete (club) fungi are known to biosynthesize an array of mostly lipophilic acetylenic natural products (Figure 10). Although acetylenic natural products have also been isolated from plants, acetylenic fatty acids rarely accumulate to substantial levels in green tissue. In contrast, it is more common to see the accumulation of acetylenic fatty acids in fungi [34]. The acetylenic fatty acid dehydrocrepenynic acid (18:2$^{\Delta 9c,12a,14c}$) has been shown to accumulate to between 60-70% of the fatty acyl content in the ectomycorrhyzial fungi of the family Cantharellaceae [42, 43].

A previous study on the fungus *Cantharellus formosus* was able to functionally characterize two putative FAD2s through heterologous expression in *S. cerevisiae* [30]. The first gene target was identified as a Δ12-desaturase (*CfDES*), producing 16:2$^{\Delta 9c,12c}$ and 18:2$^{\Delta 9c,12c}$ as 7.6% and 18.9% of the total fatty acids, respectively. The second of the putative FAD2s was revealed to accumulate 0.52% 18:1$^{\Delta 9c,12a}$ upon supplementation of linoleate, allowing this gene to be designated as a Δ12-acetylenase (*CfACET*). Upon co-expression of CfDES and CfACET, the formation of dehydrocrepenynic acid was detected, likely implicating a substrate-dependent perturbation of the Δ12-desaturase activity as being responsible for the formation of 18:2$^{\Delta 9c,12a,14c}$ in *C. formosus*. This study is the first and only time a fungal acetylenase has been isolated, to date, and provides the first enzymatic evidence for the involvement of the crepenynic acid pathway in fungal acetylenic natural product biosynthesis.

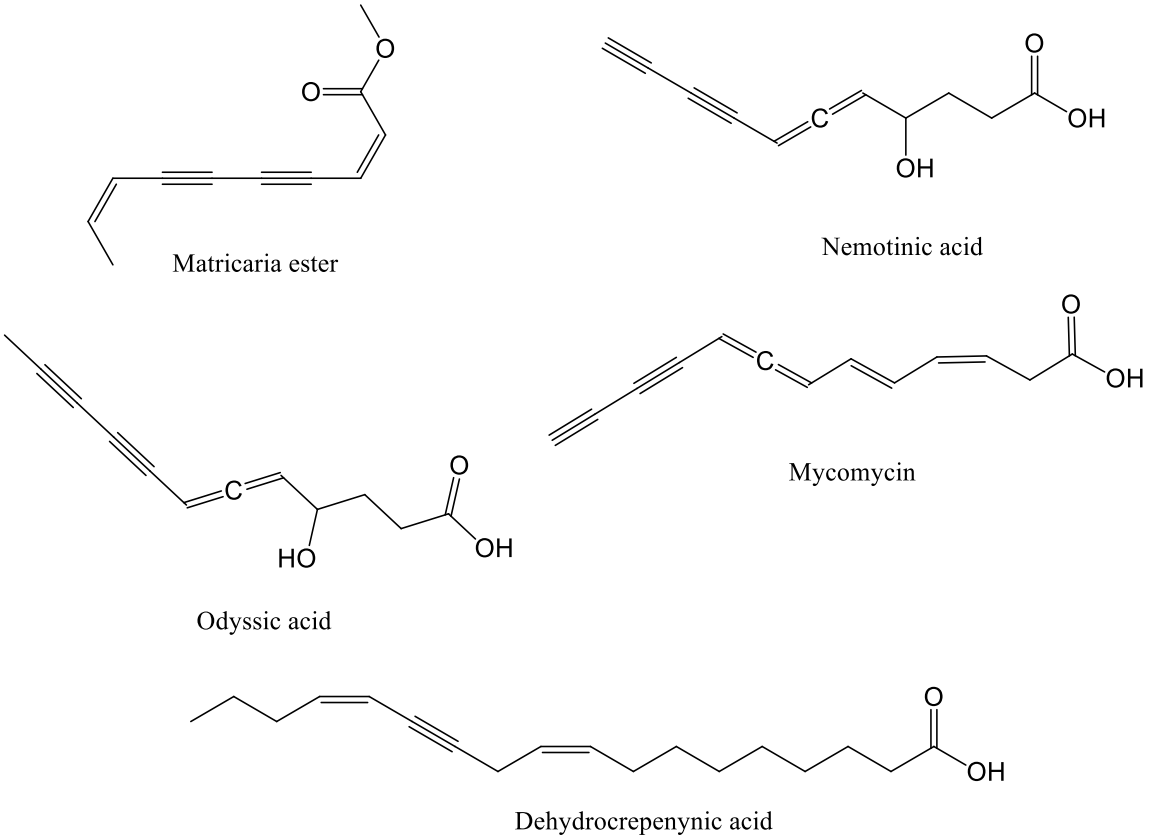


Figure 10: Examples of acetylenic natural products isolated from fungal species.

Another polyalkyne-producing basidiomycete is *Fistulina hepatica*. This edible mushroom is commonly referred to as beefsteak fungus, a moniker that is derived from its similarities in appearance of the fungus to that of raw meat. *Fistulina* is a brown-rot fungus that can typically be found living on the wood of oak and chestnut trees and has been investigated for its phenolic and volatile compounds [44, 45]. However, what sparked these studies was the isolation by Jones of five polyacetylenic natural products from liquid cultures of *F. hepatica* (Figure 11) [46]. The tetrayne-tetraol (***5***) was determined to possess modest antibacterial activity against *Staphylococcus aureus* and *Salmonella typhi* [46]. Since this initial study, other examples of bioactive polyacetylenes have been isolated from *F. hepatica* [47, 48].

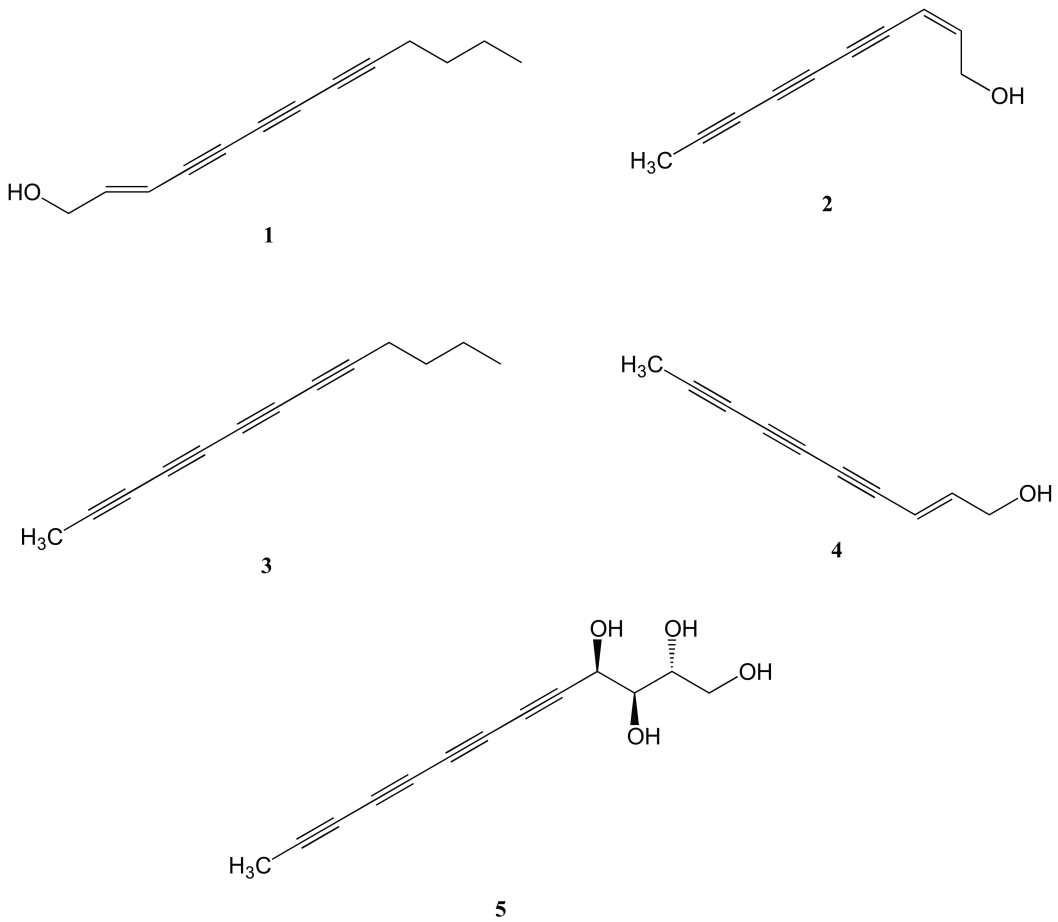


Figure 11: Polyacetylenes identified in *Fistulina hepatica* by Jones [46].

Radiolabeled tracer studies with oleic, linoleic, crepenynic, and dehydrocrepenynic acids have validated their incorporation into fungal polyacetylenes, including dehydromatricaria acid, odyssic acid, and nemotinic acid [40, 49, 50]. Taken together with the array of acetylenic metabolites produced by beefsteak fungus and the recent identification of the first crepenynate-producing fungal acetylenase from *Cantharellus*, it is hypothesized that the crepenynic acid pathway is active in *F. hepatica.* In this chapter, we investigate the involvement of the crepenynic acid pathway in the formation of alkyne-containing natural products in *F. hepatica*. This chapter describes the functional characterization of two fungal diverged desaturases: a Δ12-acetylenase and the initial example of a multifunctional fatty acid diacetylenase, capable of Δ14-desaturation/acetylenation and Δ16-desaturation. The collected evidence strongly

supports that this pathway does indeed contribute to polyacetylenic natural product synthesis in *Fistulina.* This work therefore provides the first evidence of enzymatic participants involved in the formation of intermediates beyond dehydrocrepenynate en route to polyacetylenes and supports the original crepenynic acid pathway outlined by Bu'Lock [40].

2.2 Materials and Methods

2.2.1 Preparation of Genomic DNA Library for *Fistulina hepatica*

F. hepatica (ATCC strain 64428) was grown on solid YM media at 22 °C-26 °C to reach confluence (approximately 6 weeks). The isolation of gDNA was adapted from the general method of Möller *et al.*, and included the optional RNase A treatment step [51]. Using a chilled mortar and pestle, 1.3 g of fresh *Fistulina* mycelia and fruiting bodies were frozen in liquid N_2 and ground to a fine powder in the presence of an equal amount of glass beads. The large-scale extraction of gDNA was performed as described by Möller with the volumes increased approximately 100 fold. The isolated gDNA was further purified by using Qiagen genomic tip-100 (Qiagen, Valencia, CA). Using the purified *F. hepatica* gDNA a 21.8-Mb Roche 454 shotgun sequence library was assembled by the Center for Genomics and Bioinformatics (Indiana University, Bloomington, IN).

2.2.2 Cloning and Sequencing of Putative Desaturases from *Fistulina*

Total RNA was isolated from *Fistulina* mycelia and fruiting bodies using TRIzol (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. The poly $(A)^+$ RNA was isolated from total RNA using the Dynabeads mRNA Purification Kit (Invitrogen) and was converted to rapid amplification of cDNA ends (RACE)-ready cDNA's with the SMARTer RACE cDNA Amplification Kit (Clontech, Mountain View, CA). To determine the full extent of the open-reading frames (ORFs), 5'- and 3'- gene specific RACE primers

were used in a RACE approach to obtain partial cDNA's for both *Fhe04128* and *Fhe10752* genes (Table 14, Appendix A). Using the RACE-ready cDNA's and the Advantage 2 PCR kit (Clontech), 5'- and 3'- RACE products were obtained and T/A cloned into pCR2.1-TOPO vector (Invitrogen). The RACE constructs were then transformed into DH5α chemically competent *E. coli*. Insert-containing plasmids were purified with Wizard DNA spin columns (Promega, Madison, WI) and bidirectionally sequenced.

The collected data was used to develop full-length primers containing a His_6 tag and 5'-*BamH*I and 3'-*Not*I restriction sites (Table 14, Appendix A). The full-length transcripts were obtained by Vent PCR and cloned into pYES2 expression vector using the aforementioned restriction sites and T4 DNA ligase (Gibco BRL, San Francisco, CA). The full-length, His_6 fusion- tagged constructs were transformed into competent XL1-Blue *E. coli* and sequenced to verify their integrity.

2.2.3 Expression of Fhe04128 and Fhe10752

The activities of the putative desaturases Fhe04128 and Fhe10752 were characterized using the *Saccharomyces cerevisiae* strain InvSc1. Constructs of *Fhe04128* and *Fhe10752* were transformed into baker's yeast using a standard lithium acetate-polyethylene glycol method of Gietz [52] with selection on complete minimal medium lacking uracil and containing dextrose (CM-ura+dex) for 3 days at 30 °C [53]. YPD liquid starter cultures (1% yeast extract, 2% peptone, and 2% dextrose) of single transformants were grown at 30 °C overnight with 240 rpm shaking. For expression cultures, 50 µl of fresh YPD starter culture was inoculated into 5-mL CM-ura+2% galactose and grown at 22 °C with 240 rpm for 24 h followed by expression at 15 °C with 240 rpm for 72 h. When exogenous fatty acids were fed, the 5-mL expression culture contained 0.1% (w/v) Tergitol NP-40 and the fed fatty acid at a final concentration of 0.5 mM.

2.2.4 GC/MS Analysis of Yeast Lipids

Fatty acid methyl ester (FAME), pyrrolidide (FAPy), and dimethyloxazoline (DMOX) derivatives from yeast cultures were prepared as previously described [54, 55]. In preparation for DMOX synthesis, the dried FAMEs were initially reacted with 0.5 M NaOH in 90% aqueous MeOH at 80 °C for 1 h, followed by acidification with 1 M HCl to generate the free fatty acids. The free fatty acid mixture was diluted with 1 mL of water and extracted with 2 mL of 1 : 1 hexane : diethyl ether. The organic layer containing the fatty acids was dried down under a stream of N_2 and the concentrated sample was used for DMOX derivatization. However, as derivatization of higher unsaturated fatty acids was performed, this step was carried out for 15-18 h at room temperature.

The prepared FAMEs were separated and analyzed by an Agilent Technologies GC-MS (7890A GC/ 5975C MS) using a VF-23 column (30 m x 250 µm x 0.25 µm). For "FAMEshort", the oven program was: 60 °C starting temperature, ramp 10 °C/min to 150 °C , hold 5 min; ramp 10 °C/min to 220 °C, hold 1 min; and ramp 50 °C/min to 250 °C, hold 2 min. Similarly, the pyrrolidides and DMOX derivatives were analyzed on the same GC-MS instrument and column. The pyrrolidide oven program "FAPy" consisted of: initial temperature of 150 °C, ramp 10 °C/min to 220 °C, holding 5 min; and ramp 50 °C/min to 250 °C, hold 20 min. The "DMOX" oven program was the same program used to analyze FAMEs with the exception that the hold time for the 250 °C step was 12 min. An alternate FAME program, "FAMElong", used to analyze the $18{:}2^{\Delta 9c,12a,14c}$ supplemented expressions was: 60 °C starting temperature, ramp 10 °C/min to 150 °C, hold 5 min; ramp 5 °C/min to 180 °C, hold 5 min; ramp 1 °C/min to 210 °C, hold 5 min; and ramp 10 °C/min to 250 °C, hold 3 min. Although, in all cases, the injector was set to 250 °C with a gas flow rates of 1.9 mL/min.

2.2.5 Isolation and NMR Analysis of Fatty Acid Natural Products

2.2.5.1 Purification of Natural Products

FAMEs were purified from 7 L of culture, supplemented with 0.3 mM crepenynic acid. The extracted FAMEs were purified using flash column chromatography followed by reversed-phase thin layer chromatography. The flash column was loaded with a 5.1 cm (diameter) x 17.8 cm (length) pad of SiO_2 covered by a 2.5-cm layer of sand. Initially the column was equilibrated with hexane. The mobile phase consisted of a step-down gradient from 100:1 hexane : ethyl acetate (300 mL) to 50:1 (200 mL) with an intermediate step of 75:1 (150 mL). A set of 2-mL fractions was collected over 50 tubes. This method was used predominantly to remove the saturated fatty acid methyl esters. The fractions containing mainly unsaturated fatty acids were pooled together and concentrated under reduced pressure.

Reversed-phase TLC was carried out using a method adapted from Marquardt and Wilson [56]. The pooled FAMEs from column chromatography were loaded on a Whatman KC18 reversed-phase TLC plate (20 cm x 20 cm x 200 µm). The plates were developed using 10 % (w/v) $AgNO_3$ in acetonitrile : 1,4-dioxane : acetic acid (80:20:1) at room temperature for 1 h. After air-drying, the plates were sprayed with 0.03% (w/v) 1,6- diphenyl-2,3,5-hexatriene in chloroform, air-dried and FAMEs visualized under long range UV light. The R_f values were determined and the bands were excised from the silica plate. The FAMEs were extracted from the excised silica bands using 4 mL of chloroform and the solvent was removed using a rotary evaporator. Samples were redissolved in hexanes and analyzed by GC-MS.

2.2.5.2 Structural Determination of Natural Products by NMR

Proton (500 MHz), gsCOSY, and gsHSQC NMR spectroscopic data was collected for bands corresponding to a diacetylenic and dehydrodiacetylenic FAME mixture with a

500-MHz Bruker Avance-III NMR spectrometer. The data was recorded with standard pulse sequences in $CDCl_3$ at 298 K with the chemical shifts reported as ppm referenced to residual solvent at δ 7.26 ppm.

2.2.6 Bioinformatics

Transmembrane predictions were performed using the "TopPred 0.01" program from the Mobyle server (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::toppred). The multisequence alignments and phylogenetic analysis were performed as previously described by Minto [54]. Percent amino acid identity was determined using the sequence distance function of MegAlign™ from DNASTAR Inc. (Madison, WI).

2.3 Results

2.3.1 Isolation and Cloning of *Fistulina* Sequences

Using the known desaturase and acetylenase sequences of *Cantharellus formosus* and *Phanerochaete chrysosporium,* a BlastX search through our 21.8-Mb Roche 454 shotgun sequence database yielded a list of gDNA contigs for *Fistulina hepatica*. Starting with the gDNA sequence of the resultant possible acetylenases, a BlastX search was performed to identify gene homologs from other organisms and the search results were used to hand-parse the gDNA contigs of *F. hepatic*a. The introns of the gDNA sequence were verified as the non-conserved portions of the contigs (by comparison to known desaturases) and through pattern recognition of the 5'- GTRNGT and 3'- YAG splice motifs [54].

This series of steps was performed for two *Fistulina FAD2* homologs, *Fhe04128* and *Fhe10752*. Once introns were delineated and removed, a mass alignment of the edited contigs and 13 homologs from GenBank™ was prepared to aid in the generation of the 5'- and 3'- RACE primers. The subsequent sequencing of the RACE products

verified the extent of full-length transcripts as indicated by the presence of a stop codon preceding the ATG start codon from the anticipated *fad2* sequence alignment at the 5' end of the ORF and stop codon at the 3' end of the ORF followed by the poly $(A)^+$ tail for each homolog.

Through prior work, the use of an N-terminal His_6 tag has been found to increase activity in comparison to untagged proteins when expressing diverged desaturases [30, 54]. Therefore, the collected RACE data was used to develop full-length primers containing *BamH*I-ATG-$(CAT)_6$ (sense) and *Not*I (antisense) restriction sites. The full-length transcripts were obtained by Vent PCR and cloned into pYES2 expression vector using the aforementioned restriction sites. Bidirectional sequencing was performed on pYES2 constructs to ensure sequence integrity.

Figure 12: Schematic for gene constructs. The putative desaturase is flanked by either a *BamH*I-ATG-$(CAT)_6$ fusion tag at the 5' end of the gene, or a *Not*I restriction site following the 3' stop codon.

2.3.2 Bioinformatics

The putative FAD2 cDNA's for *Fhe04128* and *Fhe10752* contained 1257 and 1242 nucleotides coding for 418 and 413 aa, respectively. Given that the *F. hepatica* mycelium was from a dikaryon, two alleles for both genes were identified and cloned. Analysis of the nucleic acid sequences for the two Fhe04128 alleles indicated the presence of 16 SNPs, while only three were identified in the Fhe10752 alleles. Comparison of the *Fhe04128* and *Fhe1075* DNA sequences revealed 54% sequence identity. In contrast to their intron-free desaturase and diverged desaturase counterparts in plants, the putative fungal desaturases, *Fhe04128* and *Fhe10752* possessed six and 12 introns, respectively, ranging 49-63 nt in length.

Primary sequence analysis identified two residue and four residue differences between the two alleles for Fhe10752 and Fhe04128, respectively. The deduced amino

acid sequence for Fhe10752 had approximately 70% sequence homology with the chanterelle Δ12-acetylenase, whereas Fhe04128 possessed a meager 41% sequence identity to both the Δ12-acetylenase and Fhe10752. Bioinformatic experiments indicated the presence of four putative transmembrane domains and three trademark histidine boxes (Figure 13), which are expected to participate in the binding of a non-heme di-iron cluster required for catalysis [10]. From the local alignment of the conserved His boxes, it appears that most diverged desaturases contain a glycine residue at the position immediately preceding the first His box (Blue arrow). However, this residue is found to be an alanine in the case of standard desaturases. Similarly, most diverged desaturases contain a glycine immediately following His box Ia, while glutamine is found in functionally characterized desaturases (Green arrow). The second His box has the sequence H(H/A)(R/K)HH in known desaturases, acetylenases, and Fhe10752 but Fhe04128 has the sequence HHAHH. Lastly, a similar observation to that of Blacklock was found for the third His box [30]. The sequence starting five residues prior to the initial His residue of the last His box (Black arrow) for acetylenases is (D/N)VX(H/N)X. This same sequence was found to be present in both alleles of Fhe10752 and Fhe04128.

2.3.3 GC-MS Analysis of *Fistulina* Tissue

Using ≈ 0.5 g of fresh *F. hepatica* mycelial tissue, FAMEs were prepared and analyzed by standard methods (Figure 14). The GC-MS lipid analysis of the tissue indicated the accumulation of 0.14% crepenynate ($18{:}1^{\Delta 9c,12a}$) and 0.49% dehydrocrepenynate ($18{:}2^{\Delta 9c,12a,14c}$). No evidence of polyacetylenic fatty acids was found in the mycelial lipid pool. FAMEs of *F. hepatica* fruiting bodies (≈ 0.2 g; a gracious gift from Professor D. Job at the University of Neuchâtel, Switzerland) were also prepared and analyzed, but lipid analysis of *Fistulina* fruiting bodies showed no accumulation of acetylenic fatty acids (data not shown).

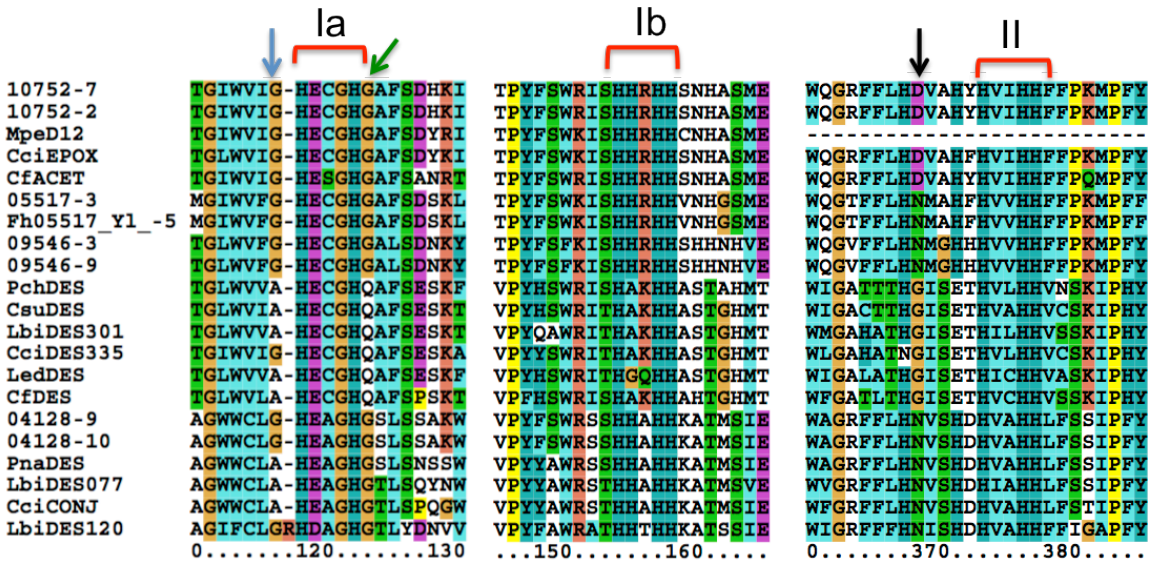


Figure 13: Local sequence alignment of the conserved His boxes for putative and functionally characterized diverged desaturases.

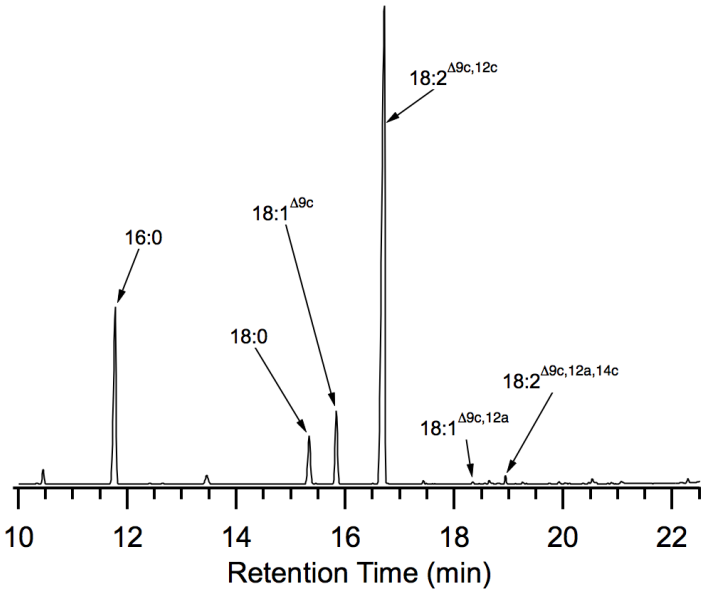


Figure 14: GC/MS total ion chromatogram of FAMEs prepared from *Fistulina hepatica* mycelial lipids. Separation of the FAMES was achieved using the "FAMEshort" oven program. FAMEs analysis indicated the accumulation of both crepenynic and dehydrocrepenynic acids to detectable levels.

2.3.4 Expression of Putative FAD2 Homologs

To functionally characterize the *Fhe04128* and *Fhe10752* genes, FAMEs of transgenic *S. cerevisiae* expressions were analyzed by GC-MS. Analysis of FAMEs developed from unfed *S. cerevisiae* expressions exhibited no new fatty acid production

by Fhe04128 but illustrated that both Fhe10752 alleles possessed Δ12 desaturase activity, as established by the presence of $16{:}2^{\Delta 9,12}$ and $18{:}2^{\Delta 9,12}$ (Figure 15). The $16{:}2^{\Delta 9,12}$ product contributed 1.79% of the total fatty acid pool while the two $18{:}2^{\Delta 9,12}$ diunsaturated fatty acids made up 1.32% of the total FAMEs, combined (Table 1). Collectively, the diunsaturated fatty acid products made up 3.11% of total FAMEs and illustrates that Fhe10752 has a regioselectivity consistent with the FAD2 desaturases.

Previous studies on the fungal Δ12-acetylenase CfACET have confirmed this enzyme's ability to generate $16{:}2^{\Delta 9,12t}$ and $18{:}2^{\Delta 9,12t}$ in the absence of linoleate substrate [30]. Comparison of the GC-MS traces for FAMEs isolated from unfed Fhe10752 and CfACET yeast expression cultures revealed that the $16{:}2^{\Delta 9,12}$ diunsaturated product contained double bonds solely in the *trans* configuration at the 12^{th} position. The $18{:}2^{\Delta 9,12}$ fatty acids generated by both Fhe10752 alleles occurred approximately three times as often in the *trans* configuration versus the *cis* isomer (Table 1).

To investigate potential Δ12-acetylenase activity, expressions of the *Fistulina* constructs were supplemented with the substrate $18{:}2^{\Delta 9c,12c}$. In all cases, the fed fatty acid was determined to accumulate to >30% of the total FAMEs. Lipid analysis of the *Fhe04128* construct indicated no new desaturase activity providing an identical lipid profile to the control expression. Conversely, analysis of GC-MS data for Fhe10752 indicated the presences of a newly synthesized fatty acid with a retention time of 18.3 min that was absent in empty vector (Figure 16).

This new fatty acid was discovered to accumulate up to 1.83% (Table 1), and had MS fragmentation consistent with that of crepenynic acid ($18{:}1^{\Delta 9c,12a}$). The identity of this Fhe10752-derived FA was later confirmed by comparing the retention time (18.3 min) and fragmentation pattern with an authentic methyl crepenynate standard prepared from *C. alpina* seed oil and comparison of pyrrolidide derivatives synthesized from both the Fhe10752 expression (Figure 17) and the methyl crepenynate standard (Appendix A). Additional inspection indicated that no other acetylenic natural products were formed, and there was no significant difference in activity between the two isolated alleles of Fhe10752.

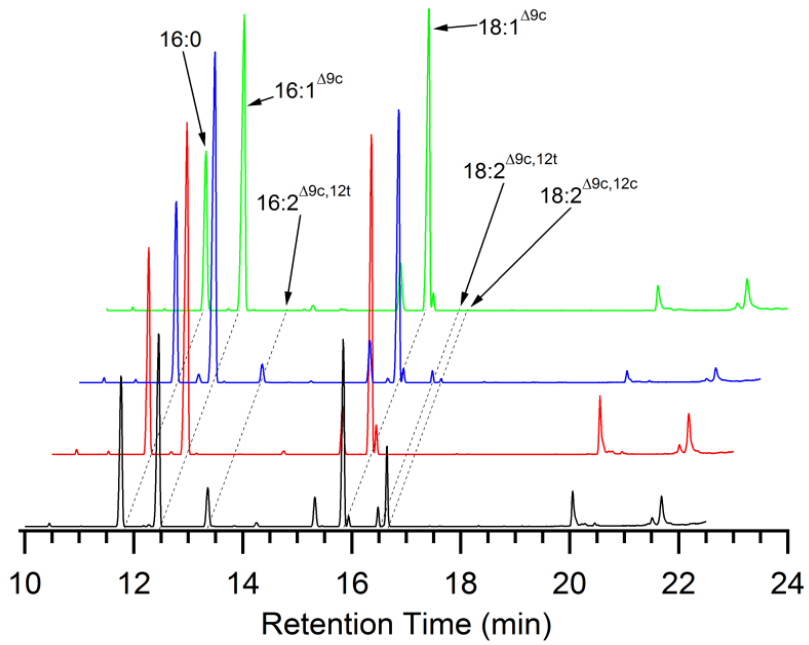


Figure 15: Demonstration of Δ12-desaturase activity. Overlay of total ion chromatographs of the total FAMEs extracted from unfed yeast expression cultures and analyzed using the 'FAMEshort" program. The traces on the left colored in black, red, blue and green represent pYES2-$(His)_6$CfACET , empty pYES2, pYES2-$(His)_6$Fhe10752, and , pYES2-$(His)_6$Fhe04128, respectively.

Table 1: Fatty acid profile for unfed and 18:2$^{\Delta 9c,12c}$-supplemented yeast expression cultures. Results are expressed as the mean % ± sd, n=3-6; *ND* = not detected.
[(a)]Fatty acid supplemented to expression cultures

Fatty Acids →	16:0	16:1 Δ^{9c}	16:2 $\Delta^{9c,12t}$	18:0	18:1 Δ^{9c}	18:1 Δ^{11c}	18:2 $\Delta^{9c,12t}$	18:2 $\Delta^{9c,12c}$	18:1 $\Delta^{9c,12a}$	26:0
Constructs										
Unfed										
pYES2	20.5±0.1	39.0±1.2	*ND*	4.55±0.04	29.8±0.5	1.89±0.05	*ND*	*ND*	*ND*	4.35±0.61
Fhe04128	18.2±0.9	36.5±0.7	*ND*	5.46±0.16	35.0±1.4	1.26±0.06	*ND*	*ND*	*ND*	3.51±1.66
Fhe10752	19.6±0.2	38.0±0.7	1.79±0.12	5.31±0.10	26.6±0.4	1.41±0.07	0.98±0.01	0.34±0.02	*ND*	5.90±0.29
+18:2Δ9c,12c [(a)]										
pYES2	18.5±0.8	24.9±0.9	*ND*	5.45±0.21	17.4±0.4	0.47±0.02	*ND*	30.7±0.4	*ND*	2.69±0.25
Fhe04128	16.6±0.4	16.1±0.8	*ND*	7.18±0.22	15.5±0.4	0.17±0.01	*ND*	41.8±0.2	*ND*	2.69±0.32
Fhe10752	19.0±0.2	21.8±0.4	1.15±0.05	5.55±0.29	15.1±0.5	0.39±0.02	*ND*	32.4±0.7	1.83±0.02	2.75±0.23

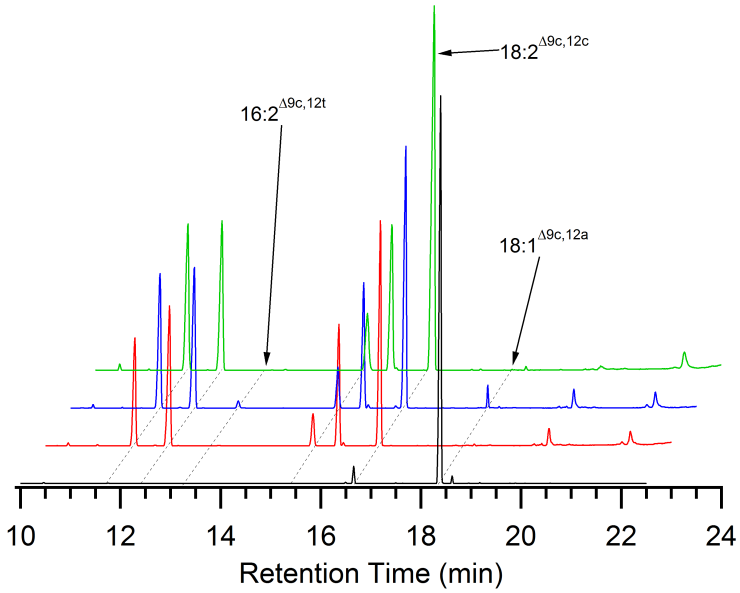


Figure 16: Demonstration of Δ12-acetylenase activity. Overlay of GC-MS analysis for total FAMEs extracted from $18{:}2^{\Delta 9c,12c}$-supplemented yeast expression cultures. The "FAMEshort" GC-MS program was used for analysis. The traces on the left colored in black, red, blue and green represent crepenynic acid, empty pYES2, pYES2-$(His)_6$Fhe10752, and pYES2-$(His)_6$Fhe04128, respectively.

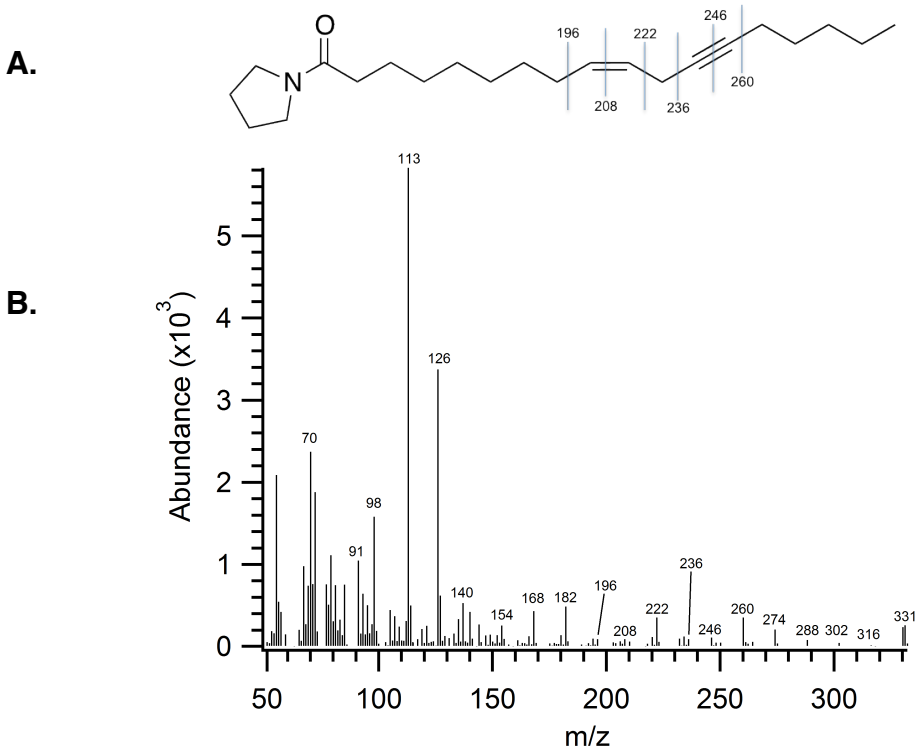


Figure 17: Analysis of mass spectra for crepenynoyl pyrrolidides. Figure (A) displays the diagnostic fragment ion sites for the crepenynic acid pyrrolidide, while figure (B) corresponds to the mass spectra synthesized from the pYES2-$(His)_6$Fhe10752 expression cultures (*m/z* 331).[1]

In order to accumulate dehydrocrepenynic acid ($18{:}2^{\Delta 9c,12a,14c}$), *F. hepatica* must contain the enzymatic machinery to desaturate the endogenous $18{:}1^{\Delta 9c,12a}$ at the C14 position, and to test this, yeast expression cultures were supplemented with 0.5 mM $18{:}1^{\Delta 9c,12a}$. The Fhe10752 construct was found to possess no new activity, demonstrating the gene's regiospecificity for desaturating C_{16}-C_{18} acyl chains at the C12 position. Interestingly, the presence of two new fatty acids was detected upon feeding $18{:}1^{\Delta 9c,12a}$ to yeast expressing Fhe04128 (Figure 18). Through comparison of relative retention times and mass spectra of authentic FAMEs and dimethyloxazoline standards isolated from *Cantharellus*, the peak at 18.9 min was identified as dehydrocrepenynic acid and was shown to accumulate to 4.59% (Table 2; Figure 19). The production of this

[1] Mass spectra for all FAME, pyrrolidine, and DMOX derivatives are located in Appendix A. Additionally, expression and standard mass spectra are stacked in the Appendix for direct comparison.

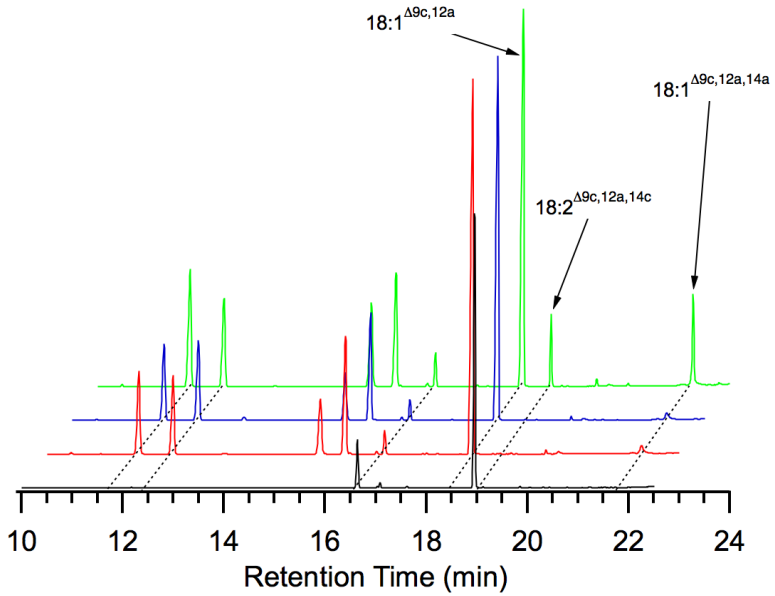


Figure 18: Demonstration of diacetylenase activity. The overlaid traces colored in black, red, blue and green correspond to total FAMES analyzed by GC-MS ("FAMEshort" program) for an authentic dehydrocrepenynate standard and yeast expression cultures of pYES2 empty vector, pYES2-$(His)_6$Fhe10752, and pYES2-$(His)_6$Fhe04128, respectively, supplemented with 0.5 mM $18:2^{\Delta 9c,12a}$.

Table 2: Fatty acid profile of $18:1^{\Delta 9c,12a}$ - and $18:2^{\Delta 9c,12a,14c}$ - supplemented yeast expression cultures. Results are expressed as the mean ± sd, n=3-6; *ND* = not detected. [(a)]Fatty acid supplemented to expression culture.

Fatty Acids →	16:0	16:1 Δ^{9c}	16:2 $\Delta^{9c,12t}$	18:0	18:1 Δ^{9c}	18:1 Δ^{11c}	18:2 $\Delta^{9c,12t}$	18:2 $\Delta^{9c,12c}$	18:1 $\Delta^{9c,12a}$	18:2 $\Delta^{9c,12a,14c}$	18:1 $\Delta^{9c,12a,14a}$	26:0	18:2 $\Delta^{9c,12a,14a,16c}$	18:2 $\Delta^{9c,12a,14a,16t}$
Constructs														
+18:1 $\Delta^{9c,12a}$ *(a)*														
pYES2	11.5±0.4	10.9±0.2	*ND*	8.60±0.37	15.6±0.6	0.10±0.02	0.20±0.02	2.18±0.08	49.1± 1.5	*ND*	*ND*	1.80±0.24	*ND*	*ND*
Fhe04128	14.0±0.5	10.4±0.4	*ND*	10.5±0.4	12.2±0.4	*ND*	0.22±0.01	2.51±0.03	38.4±1.5	4.59±0.19	7.24±0.21	*ND*	*ND*	*ND*
Fhe10752	11.6±0.3	13.0±2.3	0.49±0.06	8.33±0.22	16.2±0.7	*ND*	0.33±0.05	2.12±0.05	46.0±3.4	*ND*	*ND*	1.84±0.19	*ND*	*ND*
+18:2 $\Delta^{9c,12a,14c}$ *(a)*														
pYES2	13.6±0.4	33.5±1.5	*ND*	6.11±0.44	30.8±1.3	0.61±0.05	*ND*	2.23±0.18	*ND*	9.23±1.00	*ND*	3.87±0.46	*ND*	*ND*
Fhe04128	18.6±1.0	31.8±0.5	*ND*	6.51±0.45	21.9±0.2	0.21±0.19	*ND*	4.87±0.30	*ND*	5.42±0.41	3.91±0.48	5.43±0.30	0.97±0.04	0.39±0.04
Fhe10752	13.6 ±0.8	33.3±1.0	1.75±0.20	5.70±0.46	26.8±1.5	0.48±0.06	0.57±0.07	2.80±0.12	0.35±0.02	11.5±1.8	*ND*	3.22±0.04	*ND*	*ND*

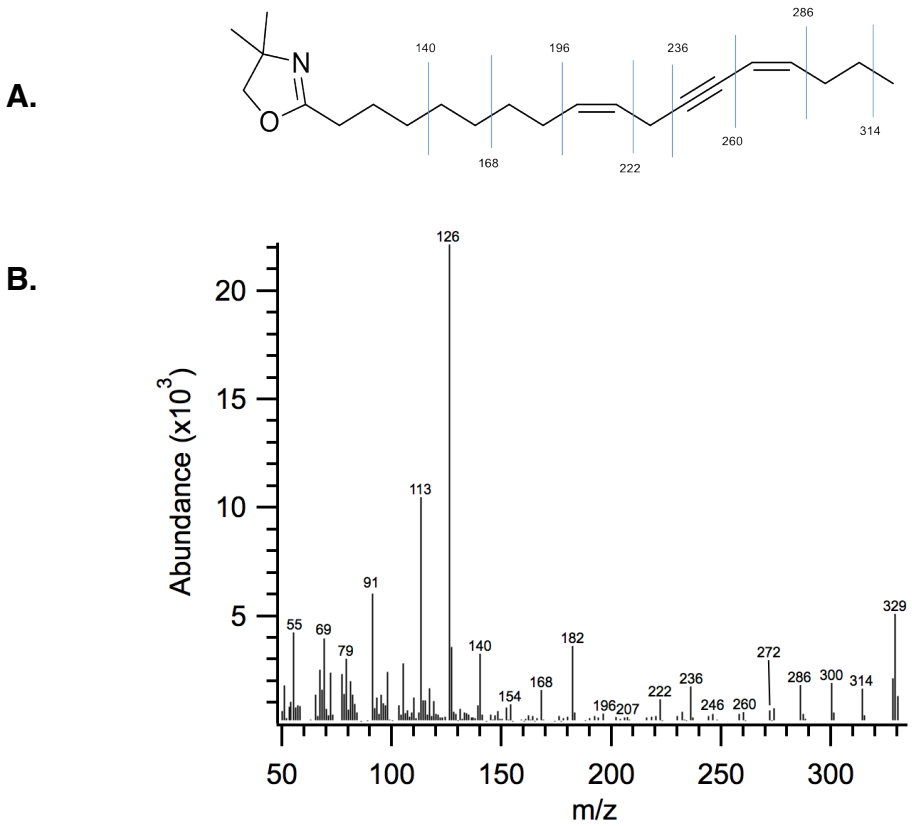


Figure 19: Mass spectral identification of the 18:2$^{\Delta 9c,12a,14c}$ DMOX derivative prepared from a crepenynate-fed Fhe04128 expression culture. Figure (A) displays the diagnostic fragments for dehydrocrepenynate DMOX, while (B) is the compound's mass spectra. The molecular ion is *m/z* 329.

acetylenic fatty acid has previously only been shown to result from desaturation of 18:1$^{\Delta 9c,12a}$ at the C14 position by FAD2 [30]. As *S. cerevisiae* is known to lack a *FAD2* gene, any desaturation occurring on the 18:1$^{\Delta 9c,12a}$ substrate is a direct result of Fhe04128 activity.

The second fatty acid product was found to co-elute with 26:0 (21.7 min), but closer inspection of the FAME MS fragmentation revealed a parent ion of *m/z* 288. This parent ion was indicative of further dehydrogenation of 18:2$^{\Delta 9c,12a,14c}$, potentially to the diacetylene 18:1$^{\Delta 9c,12a,14a}$. Evidence supporting this was gained by comparison of relative retention times and FAMEs mass spectra with a synthetic standard of 18:1$^{\Delta 9c,12a,14a}$ synthesized by R. Minto (Appendix A). To provide irrefutable evidence for

the identity of the diacetylene, DMOX derivatives were produced from the FAMEs, which were compared to derivatized standards (Figure 20, Appendix A).

The diacetylenic product accumulated up to 7.24% (Table 2); however, some of this 7.24% may have arose from co-elution of 26:0, because the two fatty acid methyl esters had similar retention times with the GC-MS oven program used. For all crepenynic acid-supplemented expressions, 18:1$^{\Delta 9c,12a}$ accumulated above 38% of the transgenic yeast lipid pool. GC-MS analysis shows that the buildup of 18:1$^{\Delta 9c,12a}$ in Fhe04128 was 38.4% compared to 46.3% and 49.2% in Fhe10752 and empty vector control, respectively. Also, acetylenic fatty acids were shown to accumulate up to 50.2% of the total fatty acids in the Fhe04128 expressions, and of this, 23.6% is the product resulting from Fhe04128 activity.

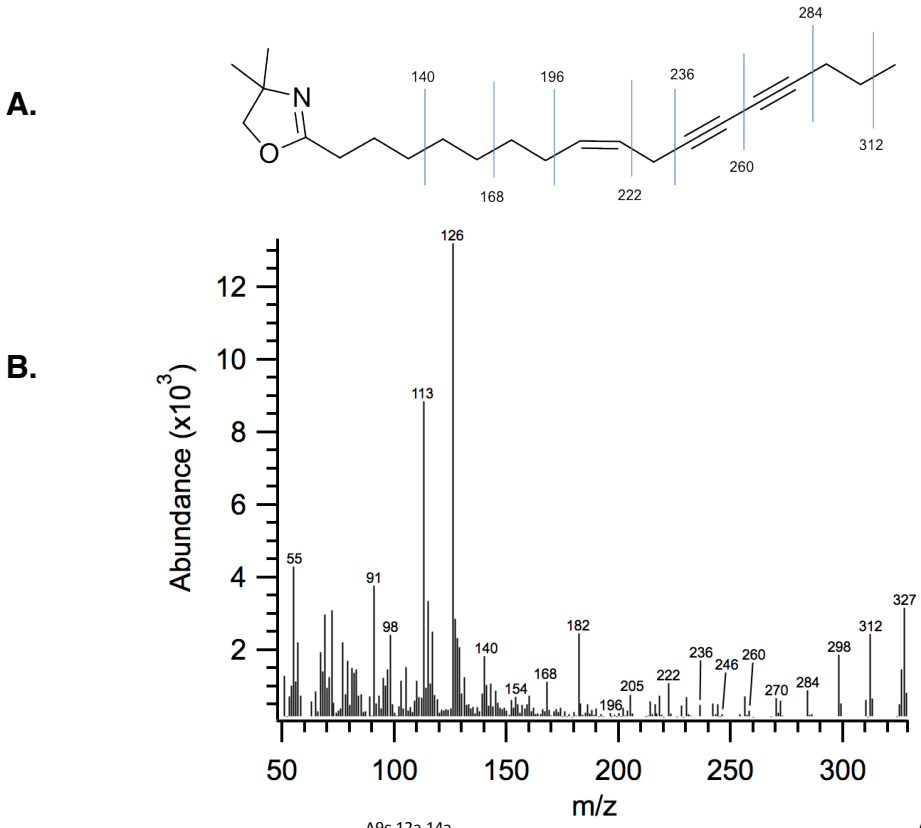


Figure 20: Mass spectrum for 18:1$^{\Delta 9c,12a,14a}$ DMOX derivative synthesized from a 18:1$^{\Delta 9c,12a}$-supplemented Fhe04128 expression culture. (A) Diagnostic fragments for diacetylenic DMOX with the exception of the molecular ion (m/*z* 327). (B) The mass spectra obtained for the analyte.

In an attempt to uncover any further activities that Fhe04128 and Fhe10752 may possess, expressions of the InvSc1 transformants with 0.5 mM-supplementation of 18:2$^{\Delta 9c,12a,14c}$ were carried out. Intriguingly, the incorporation of 18:2$^{\Delta 9c,12a,14c}$ into the heterologous yeast was found to be <12%, only 1/3 of the accumulation seen with crepenynic acid (Table 2). Also, the density of the cell cultures appeared to be lower than previous supplementation studies. Therefore, dehydrocrepenynic acid may be toxic to yeast.

For lipid analysis of the 18:2$^{\Delta 9c,12a,14c}$ supplemented expressions, the FAMEs were analyzed using a different GC-MS oven program in effort to obtain better separation between the diacetylenic natural product and 26:0. GC-MS analysis of the Fhe04128 FAMEs indicated the accumulation of dehydrocrepenynate and the diacetylene to levels of 5.42% and 3.91%, respectively (Table 2). Conversely, the accumulation of dehydrocrepenynate was almost double that of Fhe04128 in the Fhe10752 and control expression cultures. Although Fhe10752 transgenic yeast displayed no new fatty acids, FAMEs analysis of Fhe04128 also demonstrated the production of two additional fatty acids (Figure 21). The first of these peaks, with a retention time of 33.5 min, made up 0.97% of the total lipid content. The second peak at 37.3 min was found to accumulate up to 0.39%. The mass spectra of these two new fatty acids show a parent peak of *m/z* 286 (Figure 22). This is two mass units less than the diacetylene, leading to the conclusion that these two products may, in fact, be *cis* and *trans* stereoisomers of 18:2$^{\Delta 9c,12a,14a,16}$.

In an effort to conclusively identify these fatty acids, DMOX derivatives were made and analyzed by GC-MS. However, the derivatization of this highly unsaturated and conjugated fatty acid proved to be extremely challenging. The poor conversion from the FAME to the DMOX derivative, in addition to breakdown and isomerization seen in the derivatization process of this relatively unstable compound, lowered the signal-to-noise ratio in the mass spectral data and limited the interpretation of the DMOX data for the dehydrodiacetylenic natural products (Figure 23). (See discussion)

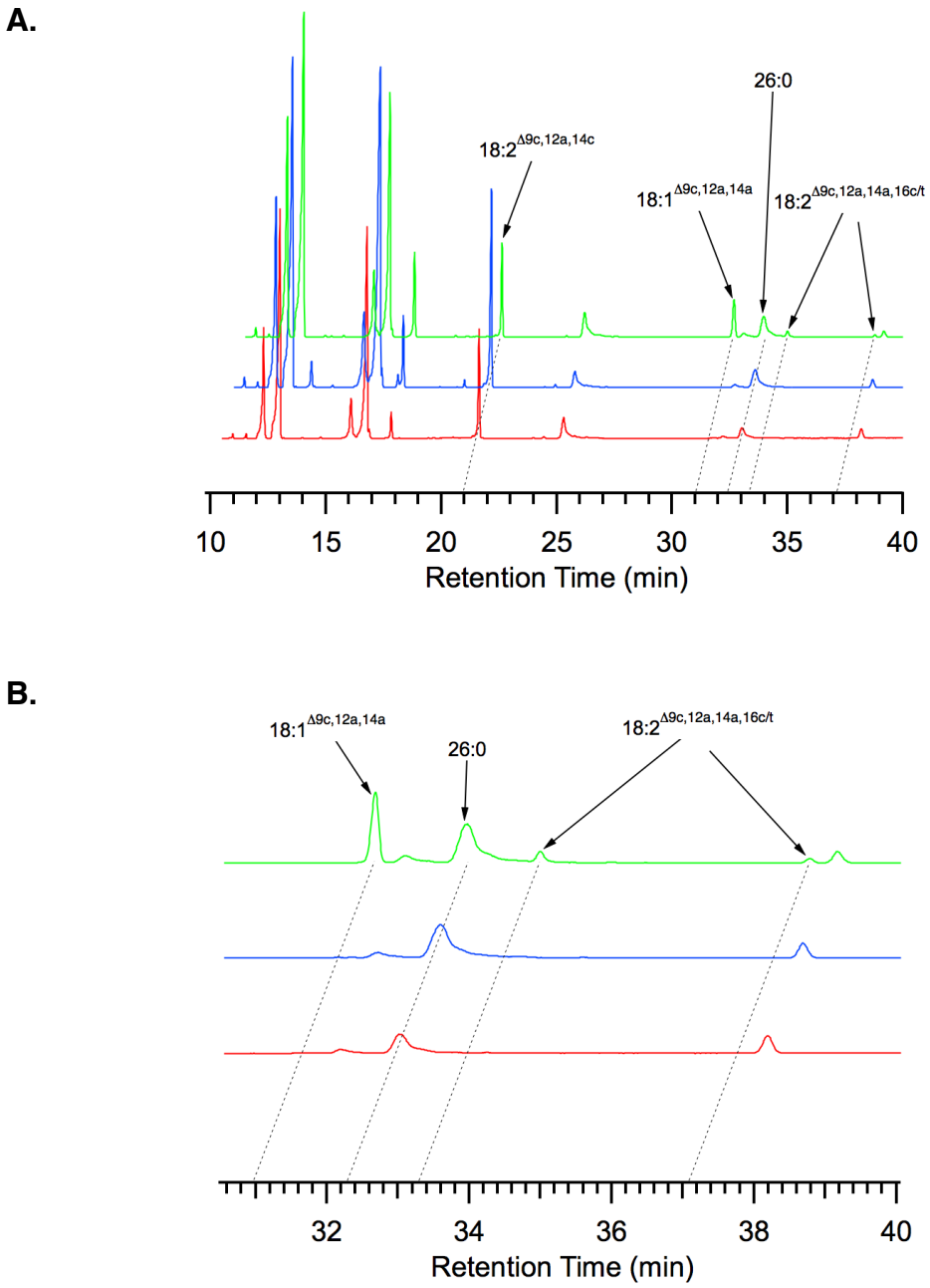


Figure 21: Demonstration of stereoselective Δ16-desaturase activity. (A) The overlaid traces colored in red, blue and green correspond to TICs of total FAMES analyzed by GC-MS, using the "FAMElong" program, for yeast expression cultures of empty pYES2, pYES2-$(His)_6$Fhe10752, and pYES2-$(His)_6$Fhe04128, respectively, supplemented with 0.5 mM $18:2^{\Delta 9c,12a,14c}$. Figure (B) is an expansion of the 30-40 min portion of (A). The peak at 37.5 min corresponds to the FAME of dioctyl phthalate.

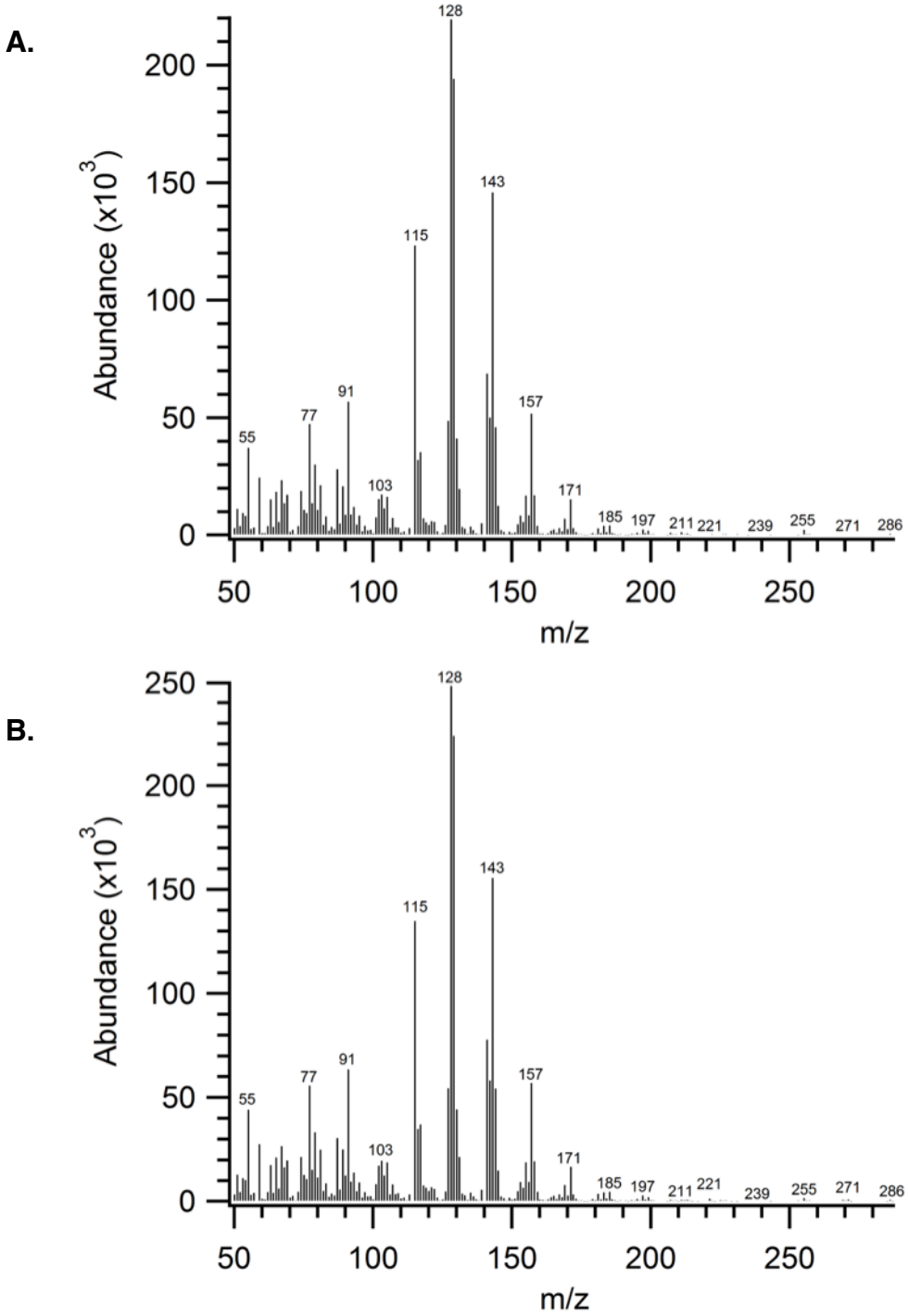


Figure 22: Mass spectrum of methyl dehydrodiacetylenes. (A) Spectra for (16*Z*)-dehydrodiacetylenic and (B) (16*E*)-dehydrodiacetylenic fatty acid methyl esters isolated from Fhe04128 transgenic yeast supplemented with dehydrocrepenynate.

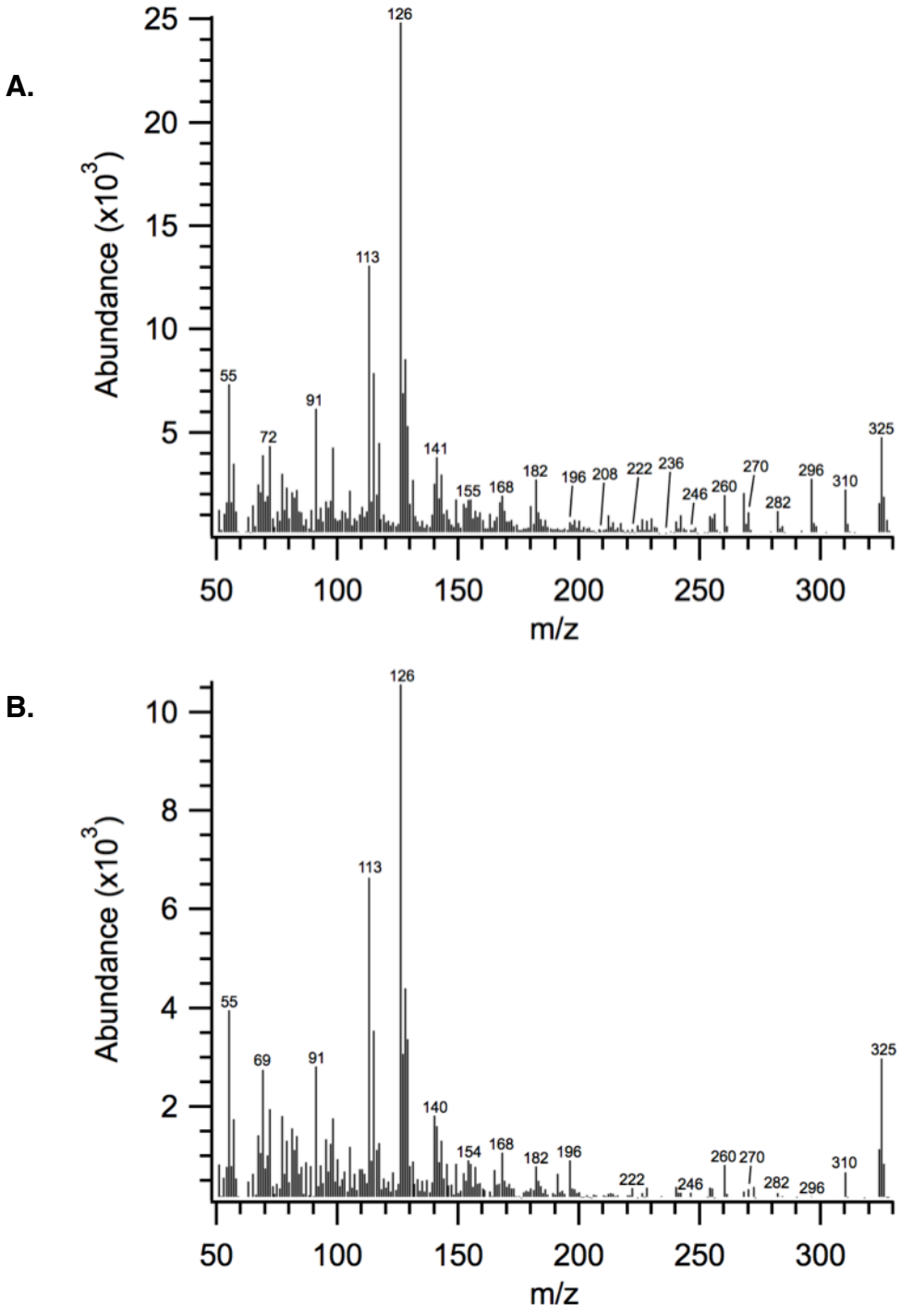


Figure 23: Mass spectra of 18:2$^{\Delta 9c,12a,14a,16}$ DMOX derivatives. Shown are the mass fragmentation patterns of (A) 18:2$^{\Delta 9c,12a,14a,16c}$ and (B) 18:2$^{\Delta 9c,12a,14a,16t}$ made from Fhe04128-expressing yeast supplemented with dehydrocrepenynate.

Furthermore, the dehydrodiacetylenic natural products are further conjugated than either C14 unsaturated product, and the last unsaturation is occurring closer to the end of the molecule. Taken together, the increase in conjugation and closer proximity of the Δ16 double bond to the methyl end has convoluted the mass spectra analysis of the dehydrodiacetylenic natural products. It appears that with increasing unsaturation toward the methyl end, the diagnostic peaks of m/z 236, 246, and 260 become less prominent. Nevertheless, many of the expected peaks corresponding to fragments for a C18 chain containing the 16-ene-12,14-diyne are present for both of the dehydrodiacetylenic species, even if at lower abundances.

2.3.5 Natural Product Purification

An alternative approach to identifying the diacetylene and dehydrodiacetylene was to isolate sufficient amounts of the natural products to use nuclear magnetic resonance (NMR) spectroscopy for structural analysis. Crepenynic acid-supplemented expression cultures were grown under the same expression conditions as were used for the screening cultures. FAMEs were generated by acidic methanolysis and isolated using a combination of flash column and reversed-phase thin-layer chromatography. After spraying the TLC plate with 0.03% (w/v) 1,6-diphenyl-2,3,5-hexatriene in chloroform and visualizing under long-wavelength UV light, the observable FAME bands were excised (Figure 24). Once the FAMEs were extracted from the silica stationary phase and analyzed by GC-MS, it was clear that the diacetylenic and dehydrodiacetylenic natural products were separated from the fatty acid pool. However, the diacetylenic and dehydrodiacetylenic natural products were incapable of being separated from each other using the current solvent system (R_f = 0.75).

Table 3: Tabulated distances for select FAMEs from reversed-phase TLC purification.

Component	R_f value
Dioctyl Phthalate	0.85
Diacetylene and Dehydrodiacetylenes	0.75
Crepenynate and Dehydrocrepenynate	0.7

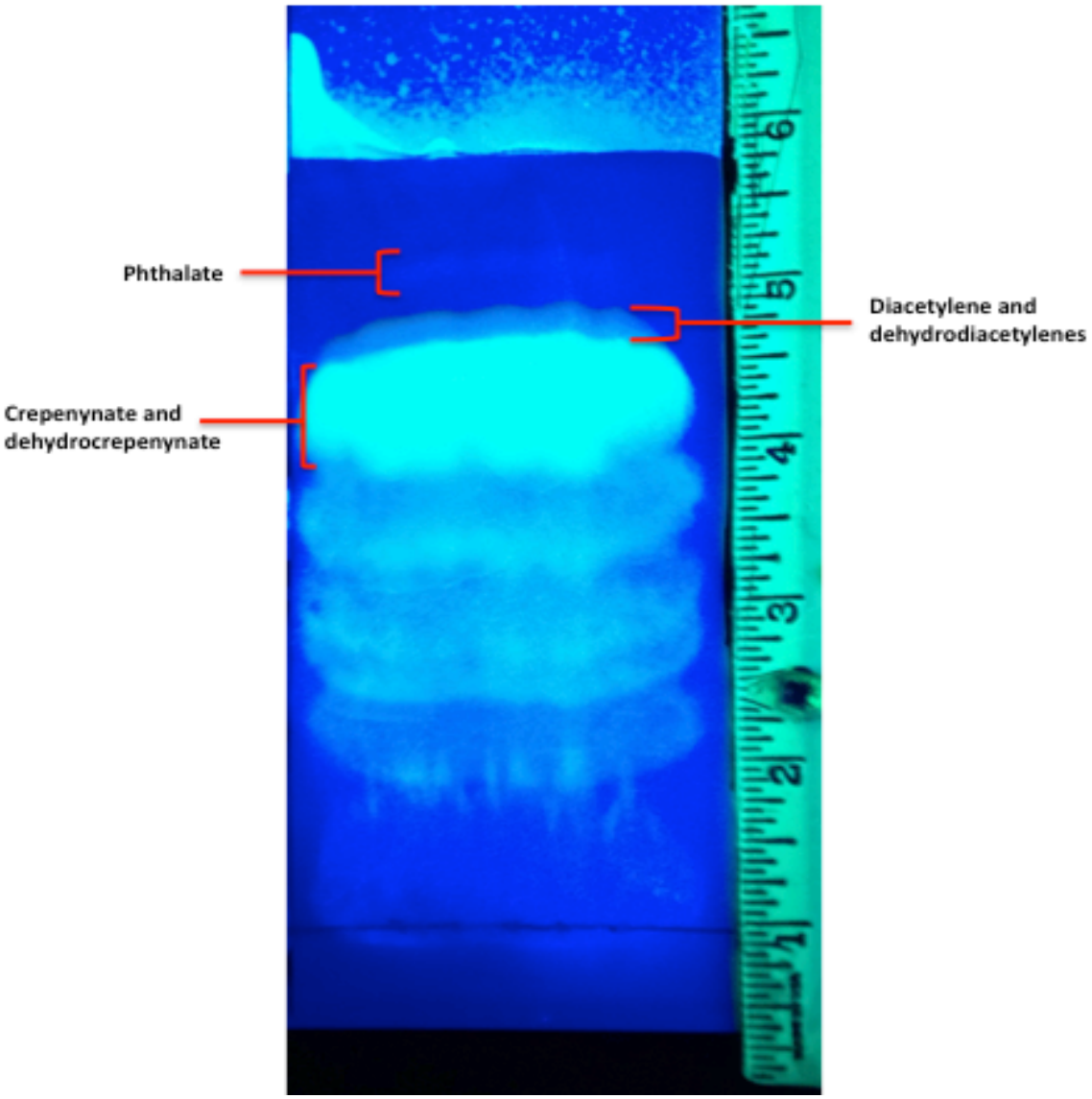


Figure 24: Reversed-phase TLC separtion of *Fistulina* natural products. Clear separation of the diacetylenic and dehydrodiacetylenic natural products was effected from the remaining FA pool using C_{18}-bonded silica plates developed in 10% (w/v) $AgNO_3$ in acetonitrile : 1,4-dioxane : acetic acid (80:20:1). The bands were excised, FAMEs extracted, and analyzed by GC-MS. Bands below the crepenynate/dehydrocrepenynate band were not analyzed.

Two subsequent purifications by reversed-phase TLC yielded a FAMEs pool containing only the diacetylenic, dehydrodiacetylenic acids, and a phthalate contaminant, as determined by GC-MS analysis. It was later concluded that TLC plates used for purification were responsible for the introduction of phthalate into the sample.

Proton NMR of the diacetylenic acid mixture indicated two minor components, each comprising approximately 1% of the sample. Proton data for the major component was consistent with the assignment of the compound being the diacetylenic acid, $18{:}1^{9c,12a,14a}$ (Table 4; Figure 25). For the minor fatty acids, vinylic multiplets at δ 6.11 and δ 6.27 and multiplets at δ 1.79 and δ 1.90 were observed by proton NMR (Figure 26). Through COSY, a correlation between the multiplet at δ 6.11 and the multiplet at δ 1.90 was observed, and is consistent with the structure of $18{:}2^{\Delta 9c,12a,14a,16c}$ (Figure 27). A second coupling was observed between protons at δ 6.27 and δ 1.79, which is consistent with $18{:}2^{\Delta 9c,12a,14a,16t}$. Coupled with the MS data, this data conclusively identifies the major acetylene as $18{:}1^{\Delta 9c,12a,14a}$ which is produced with lesser quantities of (16*Z*)- and (16*E*)- isomers of $18{:}2^{\Delta 9c,12a,14a,16}$ during Fhe04128 expression.

Table 4: Tabulated proton NMR data for diacetylenic methyl ester mixture.

Proton(s)	Multiplicity	J, Hz
2	t, 2H	7.4
8	dt, 2H	6.8, 7.1
9	m, 1H	–
10	dt, 1H	6.7, 10.4
11	d, 2H	6.7
16	t, 2H	7.0
16E/Z	m, 1H	–
17E	m, 1H	–
17Z	m, 1H	–
18	t, 3H	7.4
18E	m, 3H	–
18Z	m, 3H	–
19	s, 3H	–

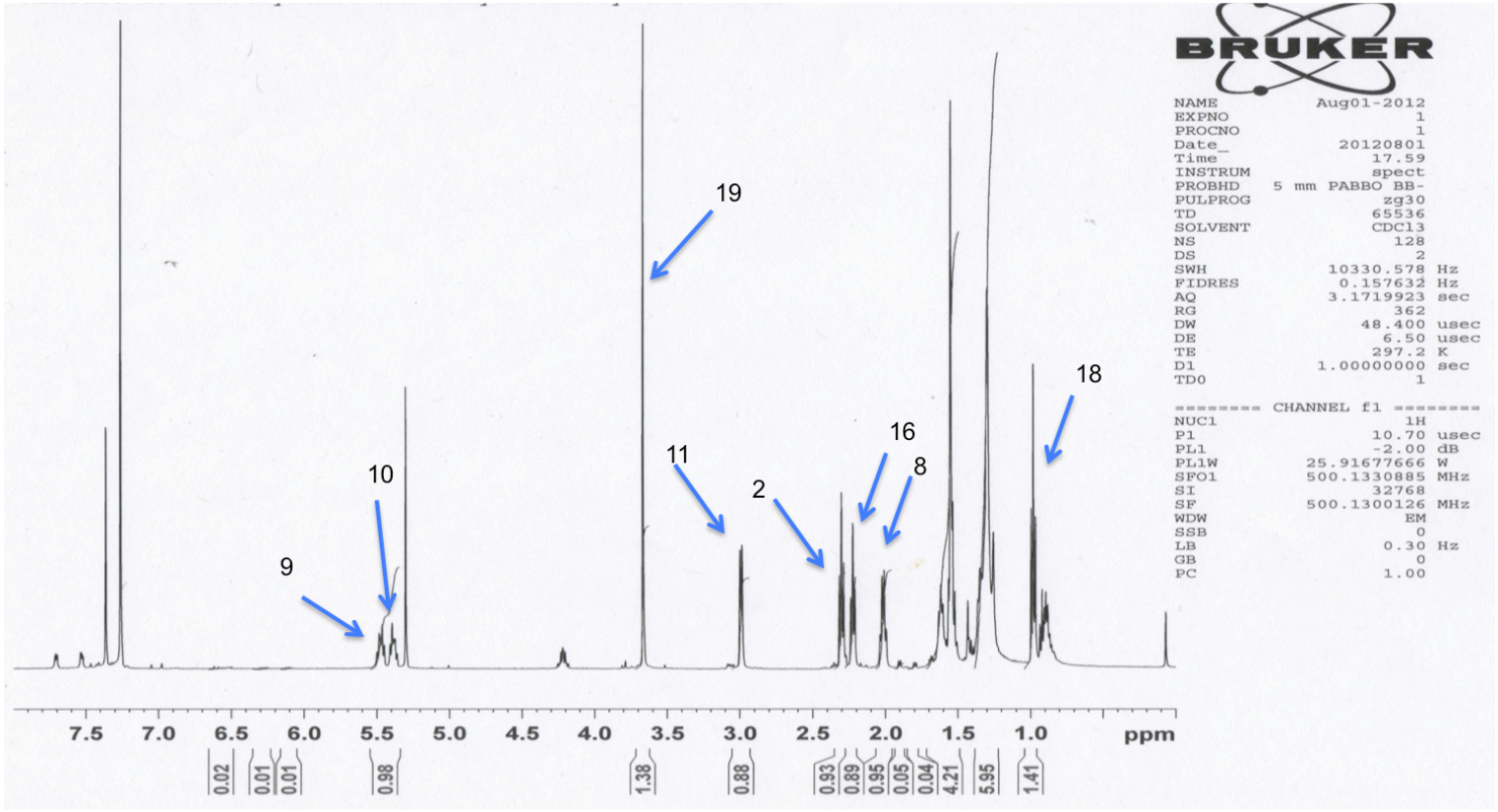


Figure 25: Structural determination of diacetylene by proton NMR.

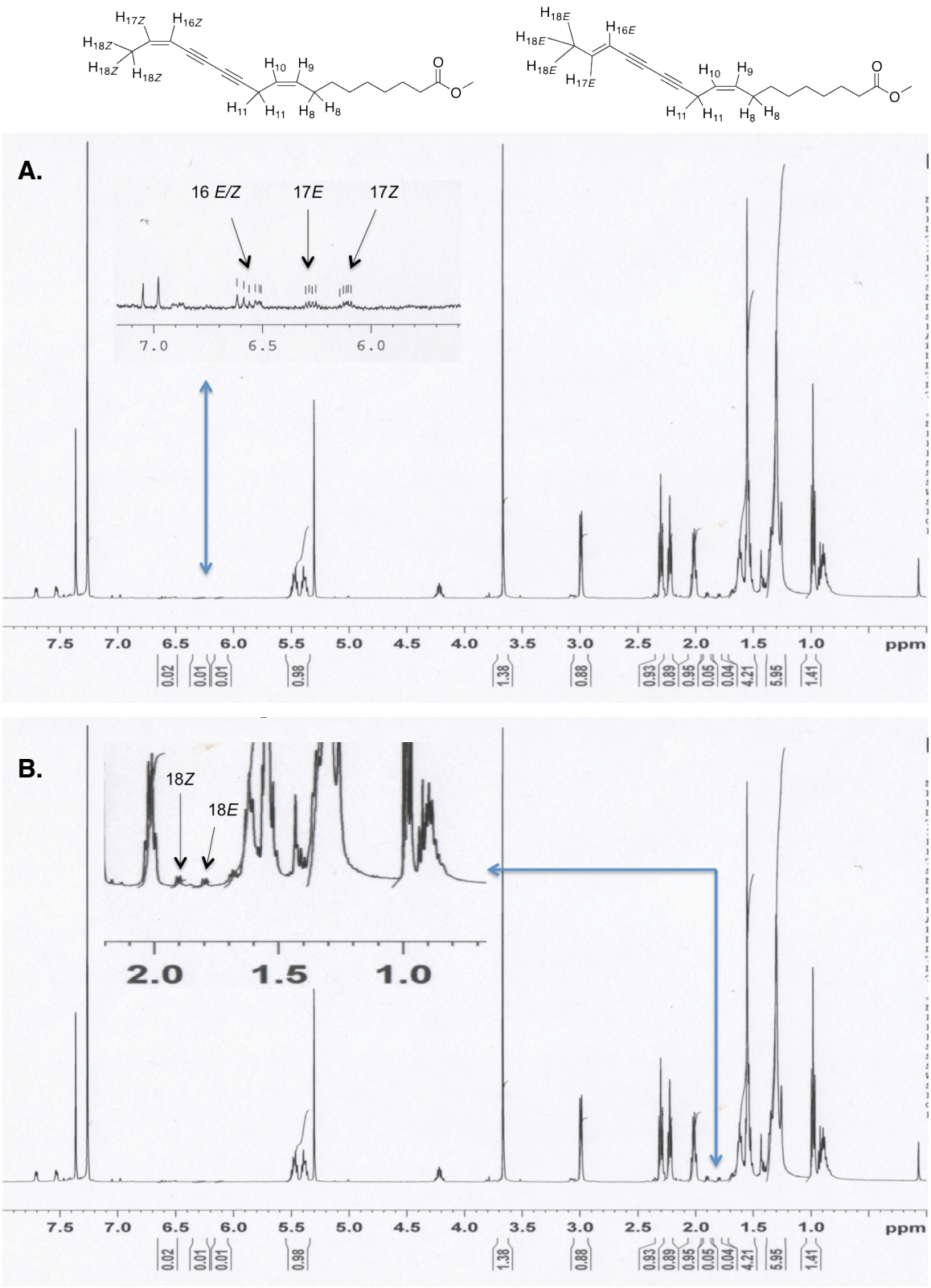
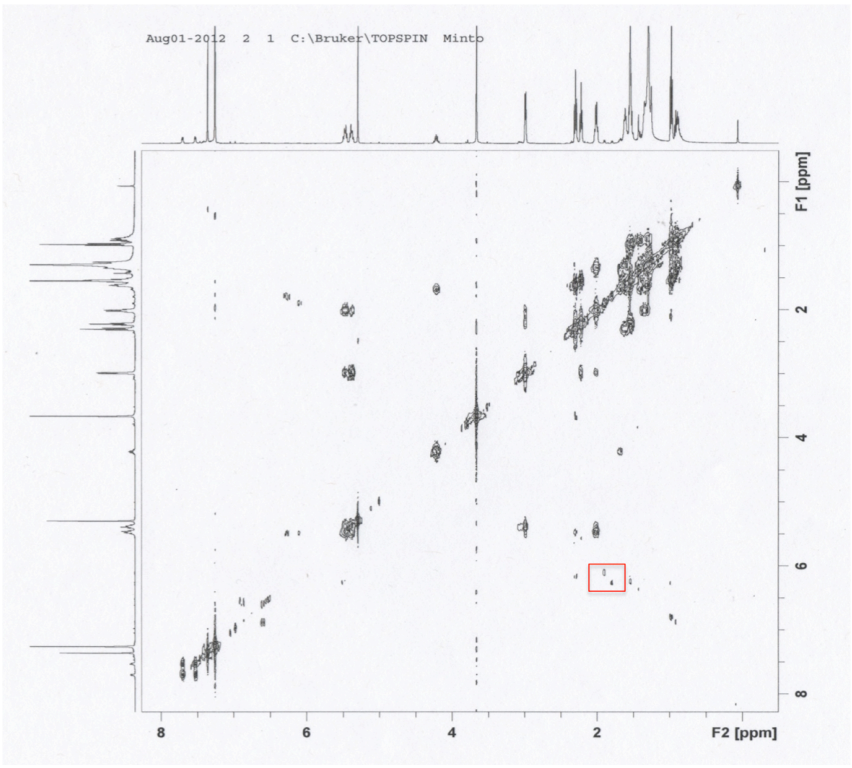


Figure 26: Structural determination of (16*Z*)- and (16*E*)- isomers of 18:2$^{\Delta 9c,12a,14a,16}$ by proton NMR.

A.



B.

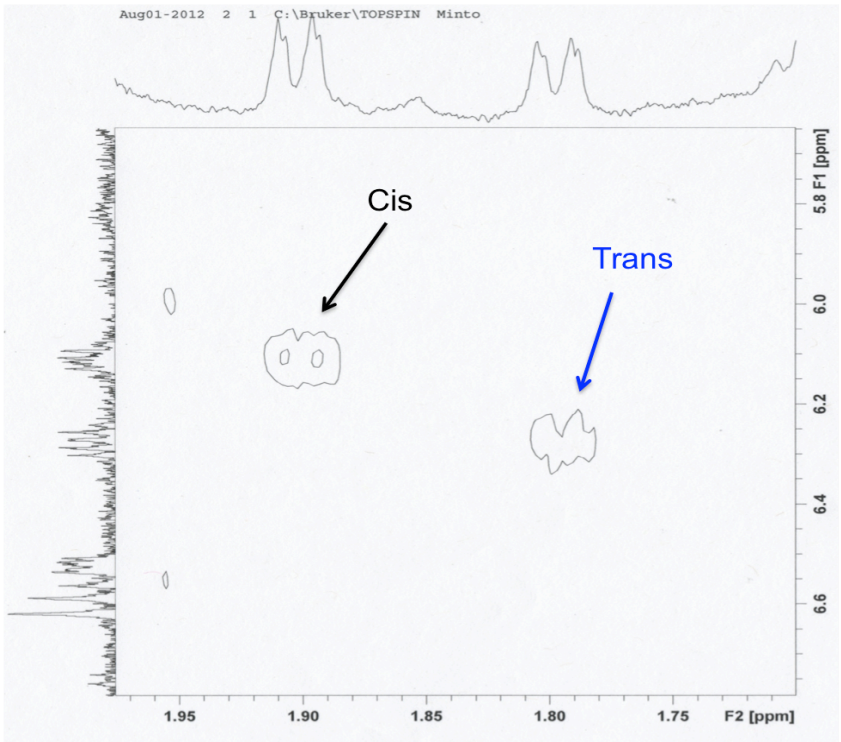


Figure 27: COSY of acetylenic natural product mixture. (A) is the full spectrum for the diacetylenic FAME mixture, which shows the identifiable correlations for the (16*Z*)- and (16*E*)- isomers of 18:2$^{\Delta 9c,12a,14a,16}$ boxed in red. (B) is a blow up of the boxed region found in (A).

2.4 Discussion

Numerous examples of acetylenic secondary metabolites have been discovered throughout the eukaryotic domain, yet the full synthetic scheme for the biosynthesis of these natural products has remained poorly defined. Based on the structures of known plant and fungal polyacetylenes and incorporation experiments with isotopically labeled compounds, the crepenynate pathway has been implicated in most cases of polyalkyne biogenesis. Because the brown-rot fungus *Fistulina hepatica* manufactures a wide array of polyacetylenic natural products [46-48], this fungus offered a clear-cut opportunity to probe the crepenynic acid pathway (Figure 28).

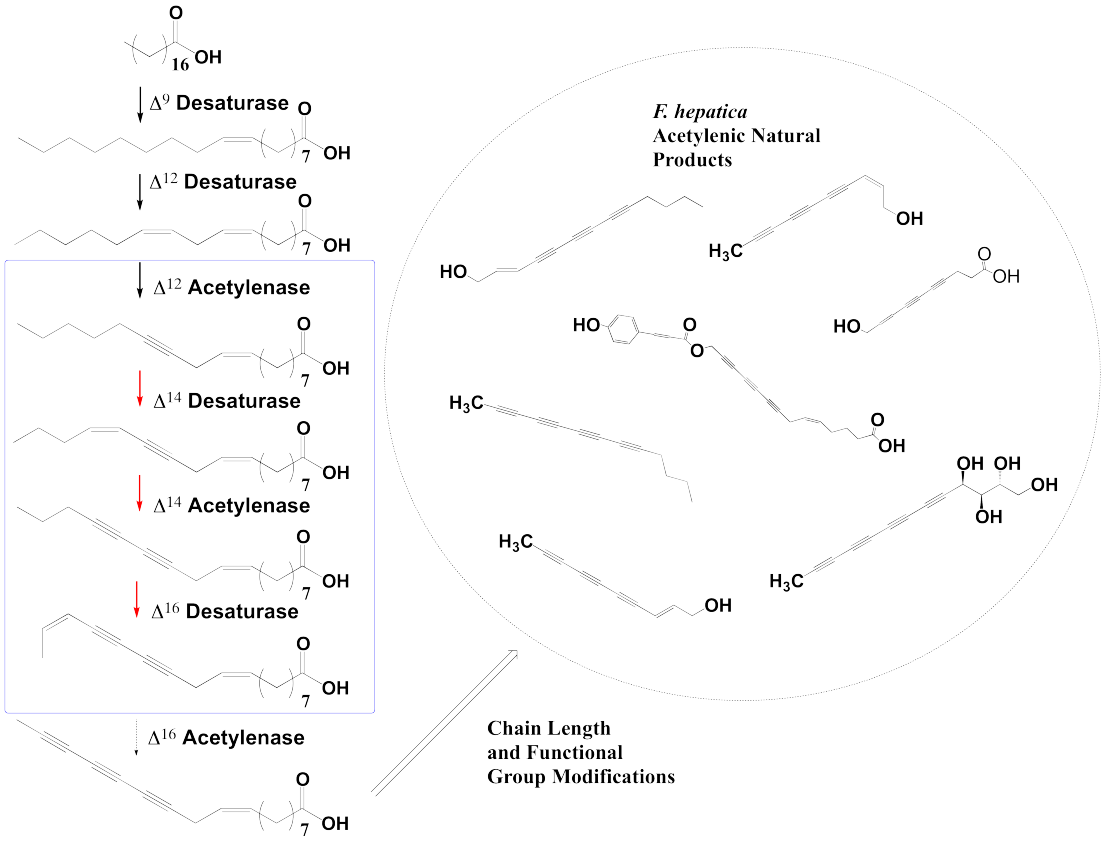


Figure 28: Proposed crepenynic acid pathway in *Fistulina hepatica*. The steps in the pathway known prior to this study are indicated by black bold arrows, steps proven through this study are marked by red arrows, and the theoretical steps are indicated by dashed arrows. Fhe04128 and Fhe10752 are capable of catalyzing the enzymatic steps contained in the blue box. The *F. hepatica* natural products were reported elsewhere [46-48].

FAMEs analysis of *F. hepatica* mycelia confirmed the accumulation of both 18:1$^{\Delta 9c,12a}$ and 18:2$^{\Delta 9c,12a,14c}$. To accumulate 18:2$^{\Delta 9c,12a,14c}$, *Fistulina* must contain at least one FAD2 Δ12-desaturase and one FAD2-like Δ12-acetylenase responsible for Δ12-/Δ14-desaturation and Δ12-acetylenation, respectively. Our investigation began with identifying FAD2 desaturase homologs within the *F. hepatica* genome. Two of the cloned putative desaturases, Fhe04128 and Fhe10752 were discovered to contain the typical characteristics of membrane-bound desaturases. Sequence analysis uncovered the presence of the three catalytically essential His boxes and four transmembrane domains that act to orient the His boxes on the cytosolic side of the membrane. The deduced amino acid sequence for Fhe10752 had substantial sequence homology to known FAD2 desaturase homologs, in particular the Δ12-acetylenase CfACET; while, the Fhe04128 polypetide was found to have significantly less sequence homology to previously characterized desaturases.

The demonstration of Δ12-desaturase activity by Fhe10752, under unfed expression conditions, supported the enzyme being placed in the FAD2 desaturase family. However, the formation of C16 diunsaturated products preferentially in the *trans* configuration at C-12, and a mixture of C18 isomers at the 12th position was reminiscent of the activity seen from characterized Δ12-acetylenases [19, 30].

The formation of 18:1$^{\Delta 9c,12a}$ upon supplementation of 18:2$^{\Delta 9c,12c}$ to Fhe10752 expressions (as validated by pyrrolidides), demonstrated this protein possesses the anticipated Δ12-acetylenase activity. A closer look at the conversion percentage [*(product / sum of direct substrate and product) x100]* for the formation of 16:2$^{\Delta 9c,12t}$ and 18:1$^{\Delta 9c,12a}$ in Fhe10752 when 18:2$^{\Delta 9c,12c}$ was supplemented demonstrated that Fhe10752 preferentially acts on the linoleate substrate to form crepenynic acid (Table 5). No activity other than the Δ12-desaturase/Δ12-acetylenase activities were displayed by Fhe10752 and is indicative of the enzyme's specificity for Δ12-dehydrogenation on 16:1$^{\Delta 9c}$, 18:1$^{\Delta 9c}$, and 18:2$^{\Delta 9c,12c}$ substrates. This multifunctional enzymatic activity displayed in Fhe10752 is not unique to Δ12-acetylenases, but rather has become a reoccurring theme with FAD2-like desaturases, as discussed in Chapter 1 and the

background section of this chapter. This is the first time a Δ12-acetylenase has been isolated and functionally characterized from *F. hepatica*, and provides support for the hypothesis that polyacetylenes located in *Fistulina* are being biosynthesized via the crepenynic acid pathway.

Table 5: Tabulated conversion ratios for Fhe04128 and Fhe10752 constructs. [A] Values are estimates due to the presence of 18:2$^{\Delta 9c,12c}$ in the crepenynic and dehydrocrepenynic acid feedstocks.

Construct	Product	Conversion Percent
Fhe10752 Unfed	16:2$^{\Delta 9c,12t}$	4.5
Fhe10752 Linoleate Feeding	16:2$^{\Delta 9c,12t}$	5.0
Fhe10752 Crepenynate Feeding	16:2$^{\Delta 9c,12t}$	3.6
Fhe10752 Dehydrocrepenynate Feeding	16:2$^{\Delta 9c,12t}$	5.0
Fhe10752 Unfed	18:2$^{\Delta 9c,12t}$	3.5
Fhe10752 Crepenynate Feeding	18:2$^{\Delta 9c,12t}$	1.8[A]
Fhe10752 Dehydrocrepenynate Feeding	18:2$^{\Delta 9c,12t}$	1.9[A]
Fhe10752 Unfed	18:2$^{\Delta 9c,12c}$	1.3
Fhe10752 Crepenynate Feeding	18:2$^{\Delta 9c,12c}$	11.4[A]
Fhe10752 Dehydrocrepenynate Feeding	18:2$^{\Delta 9c,12c}$	9.3[A]
Fhe10752 Linoleate Feeding	18:1$^{\Delta 9c,12a}$	5.3
Fhe10752 Dehydrocrepenynate Feeding	18:1$^{\Delta 9c,12a}$	11.1
Fhe04128 Crepenynate Feeding	18:2$^{\Delta 9c,12a,14c}$	9.1
Fhe04128 Crepenynate Feeding	18:1$^{\Delta 9c,12a,14a}$	14.4
Fhe04128 Dehydrocrepenynate Feeding	18:1$^{\Delta 9c,12a,14a}$	36.6
Fhe04128 Dehydrocrepenynate Feeding	18:2$^{\Delta 9c,12a,14a,16c}$	9.1
Fhe04128 Dehydrocrepenynate Feeding	18:2$^{\Delta 9c,12a,14a,16t}$	3.6

Through comparison with sequence databases, Fhe04128 was distantly related to the known FAD2 desaturases. Initial expressions of Fhe04128 in the presence or absence of exogenous linoleate did not result in detectable desaturase activity. Given that *Fistulina* is capable of producing highly unsaturated acetylenic compounds beyond crepenynic acid, it was speculated that 18:1$^{\Delta 9c,12a}$ might be the substrate required for

Fhe04128. Crepenynic acid-supplemented expressions began to reveal the unprecedented cascade of desaturation activities possessed by Fhe04128. This super-desaturase contains the catalytic capacity to carry out a multitude of desaturation steps never seen before.

Fhe04128 is a tri-functional enzyme capable of Δ16-desaturation in addition to Δ14 desaturation/acetylenation. The desaturase was able to produce dehydrocrepenynic acid, $18{:}1^{\Delta 9c,12a,14a}$, and the (16*Z*)- and (16*E*)- isomers of $18{:}1^{\Delta 9c,12a,14a,16}$ when supplemented $18{:}1^{\Delta 9c,12a}$ and $18{:}2^{\Delta 9c,12a,14c}$. Although the dehydrodiacetylenic natural products were initially only observed when $18{:}2^{\Delta 9c,12a,14c}$ was supplemented, partial purification of the Δ16-unsaturated products was achieved through reversed-phase TLC of crepenynic acid-supplemented expressions. This demonstrates the first time an acetylenase has possessed enough activity to generate its own substrate for two consecutive desaturation steps; the formation of the diacetylene from the dehydrocrepenynate product, and its use as the substrate for the formation of $18{:}2^{\Delta 9c,12a,14c,16c/t}$. However, a closer look at the conversion ratios for both $18{:}1^{\Delta 9c,12a}$ and $18{:}2^{\Delta 9c,12a,14c}$ supplementation studies demonstrates that Fhe04128 preferentially dehydrogenates dehydrocrepenynic acid to form the $18{:}1^{\Delta 9c,12a,14a}$ product (Table 5).

Although the formation of the tri-acetylene was not detected, the *cis/trans* ratio of dehydrodiacetylenic products generated by Fhe04128 was nearly identical to the *trans/cis* ratio of C18 diunsaturated products generated by Fhe10752. The unfed Fhe10752 expressions revealed that $18{:}2^{\Delta 9c,12t}$ was produced 2.88 times more frequently than $18{:}2^{\Delta 9c,12c}$. Additionally, it was found that the dehydrocrepenynate supplemented expression of Fhe04128 produced the $18{:}2^{\Delta 9c,12a,14a,16c}$ 2.49 times more frequently than the *trans* isomer. Coupled with the formation of *cis/trans* isomers found in other acetylenases, the close proximity of these ratios may suggest Δ16-acetylenase activity is within the catalytic abilities of Fhe04128, but was too low to be detected using the current expression system [19, 30, 32].

Previous expressions of Δ12-acetylenases in soybean and yeast have shown the production of 18:2$^{\Delta 9c,12a,14c}$. The origin of formation for dehydrocrepenynate was originally believed to be carried out in one of two ways [28]. The first hypothesis was crepenynate desaturation by a FAD3 (Δ15-desaturase) at the 14th position due to the triple-bonded, altered substrate fitting into the desaturase active site. The second hypothesis speculated that the Δ12-acetylenase could use α-linolenic acid (18:3$^{\Delta 9c,12c,15c}$) as a substrate to form the Δ12-acetylenic bond with concomitant rearrangement of the Δ15-double bond to the 14th position. However, a co-expression of the *Cantharellus* Δ12-desaturase and Δ12-acetylenase in yeast revealed the formation of dehydrocrepenynic acid [30]. This is consistent with an alternative theory that the Δ12-desaturase is responsible for the desaturation of crepenynic acid at the Δ14 position to form dehydrocrepenynate in *C. formosus* and is most likely the origin in other expression systems.

In contrast, Fhe04128 expressions demonstrate the accumulation of 18:2$^{\Delta 9c,12a,14c}$, but Fhe04128 indicated no Δ12-desaturase activity when expressed in the absence of exogenous fatty acids. It is also known that *Saccharomyces* lacks an oleate desaturase; hence, the formation of dehydrocrepenynic acid is a direct result of Δ14-desaturase activity exhibited by Fhe04128. Since the conversion ratios for Fhe04128 demonstrate that this enzyme preferentially acts as a Δ14-acetylenase, the larger accumulation of dehydrocrepenynic acid over 18:1$^{\Delta 9c,12a}$ in *Fistulina* mycelia may result from a combination of FAD2 and Fhe04128 activities. The lack of accumulation of Fhe04128 products beyond 18:2$^{\Delta 9c,12a,14c}$ in mycelia may demonstrate either a degradation of the dehydrocrepenynate product within the tissue, or a rapid conversion of the 18:1$^{\Delta 9c,12a,14a}$ and 18:2$^{\Delta 9c,12a,14a,16c/t}$ compounds into other natural products along the polyacetylene biosynthesis pathway.

Validation of the membrane-bound fatty acyl desaturation mechanisms has remained elusive over the years due to a lack of structural and enzymological data on purified enzymes. However, kinetic isotope effect studies have confirmed that fatty acyl chain desaturation at Δ8, Δ9, Δ11, Δ12, and Δ13 and acetylenation at the Δ11 and Δ12

all proceed by initial pro-(*R*) hydrogen abstraction at the carbon atom nearest the carboxyl end [17, 18, 26, 32, 57]. It is generally accepted that the formation of *cis* double bonds at these positions results from abstraction of the pro-(*R*) hydrogen at their respective carbon atom, followed by a rapid collapse of the unstable intermediate and ultimately resulting in the loss of the pro-(*R*) hydrogen located on the adjacent carbon [19, 30, 57]. It is theorized that the unstable intermediate formed during desaturation eliminates far to quickly for a substrate conformational change to occur within the active site. Therefore, it is conceivable that the desaturation reactions performed by Fhe04128 would be carried out with proton abstraction occurring first at the C14 position for both desaturation [pro-(*R*)] and acetylenation and at the C16 position for desaturation. However, the pro-(*R*) abstraction at C17 would most likely take place in the formation of 18:2$^{\Delta 9c,12a,14a,16c}$ and pro-(*S*) for the formation of the 18:2$^{\Delta 9c,12a,14a,16t}$ [19] (Figure 29).

The reason for switching the fatty acid natural product derivatization from pyrrolidides to oxazolines merits discussion. As the natural products became more unsaturated, greater susceptibility to derivatization problems was observed. This is consistent with several studies investigating the derivatization of other highly unsaturated fatty acids. Crepenynic acid has been shown to undergo rearrangement and isomerization in the presences of both high temperature and basic conditions, together [58]. More recently, isomerization and cyclization of *trans*-3-hexadecenoic acid and crepenynic acid, respectively, has been seen in preparation of DMOX derivatives using the one-pot method at 180 °C [55, 59].

The one-pot derivatization method for converting FAMEs, through free fatty acids, to pyrrolidides under mildly basic conditions at 100 °C was successful only for methyl crepenynate. Conversely, attempts to extend the method to the pyrrolidides of the desaturation products of Fhe04128 proved unsuccessful. It is presumed that the polyacetylenes are breaking down under the high heat and mildly basic conditions used for pyrrolidide formation. After several failed attempts, DMOX derivatives were employed for characterization of the Fhe04128 products. While DMOX preparation

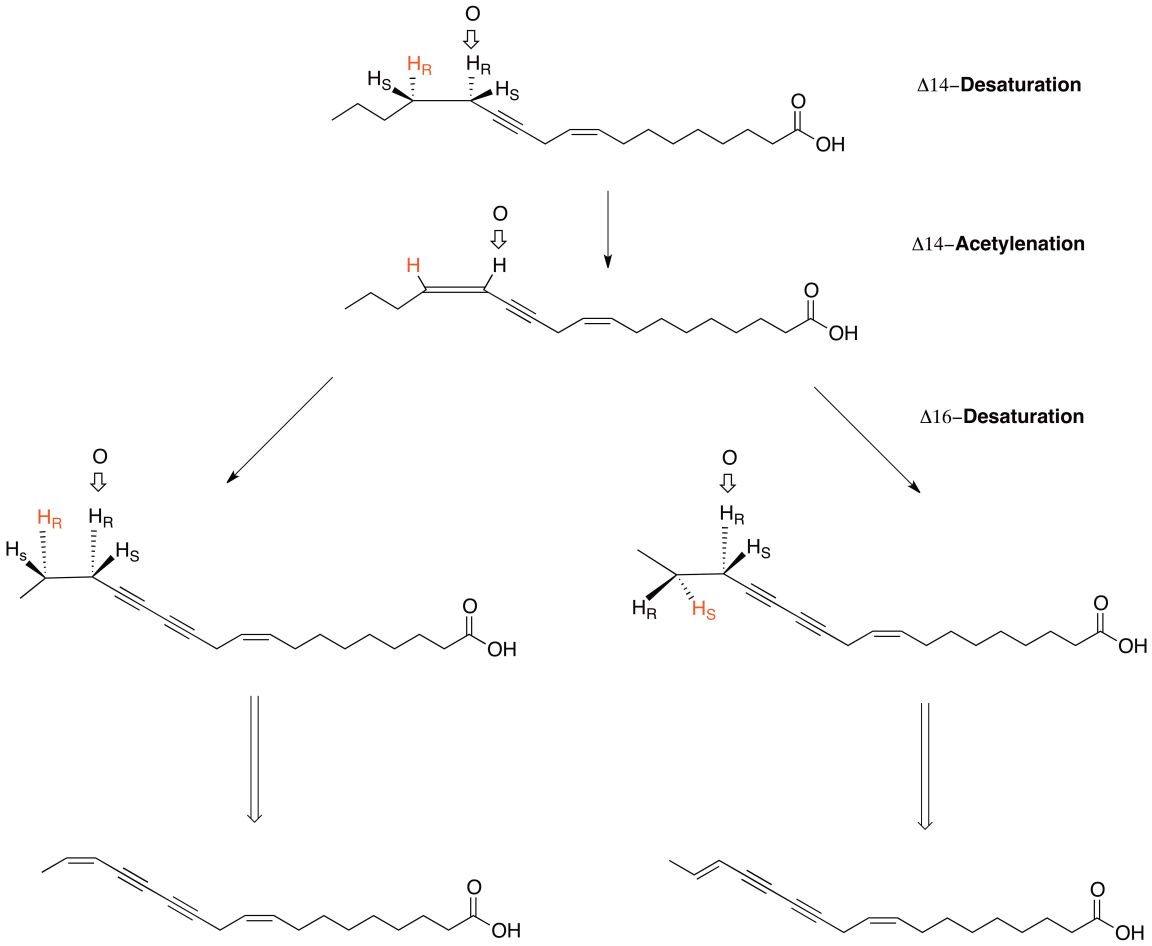


Figure 29: Predicted model for Fhe04128 dehydrogenation reactions. Experimental data has demonstrated that membrane-bound desaturation occurs with an initial removal of the pro-(*R*) hydrogen closest to the carboxyl end. It is generally accepted that the short lifespan of the highly unstable intermediate does not allow for a substrate conformational change within the active site, and the reaction would therefore proceed via *syn* stereochemistry. The hydrogen atoms labeled in red correspond to the second hydrogen atom removed during each desaturation step.

appears to be comparable in basicity to the pyrrolidide method, the ability to run the milder derivatization at a lower temperature seemed advantageous [60].

The FAMEs must be converted to their free acids before generating the acid chlorides required for DMOX derivatization. The free acids were prepared by a simple incubation of the FAMEs with 0.5 M NaOH at 80 °C for 1 h, followed by acidification with 1 M HCl. The first attempt at DMOX derivatization was successful for the quantitative

characterization of dehydrocrepenynate but turned out unfruitful for the diacetylene. Through further examination of the total ion chromatograms, it was realized that the crepenynate/dehydrocrepenynate ratio was lower in the starting FAMEs sample than the ratio in the resulting DMOX product pool. This was a clear sign of dehydrocrepenynate degradation under mildly basic conditions at high temperature and is again consistent with the literature [58-60] From this, it was anticipated that more highly unsaturated products than dehydrocrepenynate were much more susceptible to isomerization under the current conditions.

Ultimately, hydrolysis of the diacetylenic lipids were carried out using the same 0.5 M NaOH solution for conversion to the free acid as previous attempts, but at room temperature overnight as opposed to 80 °C for 1 h. This proved to be successful. For the dehydrodiacetylenic oxazolines, the conversion of the FAME to the DMOX derivative remained substantially less efficient. The inherent issues described above are compounded by the fact that the starting material for dehydrodiacetylenic natural product conversions was significantly lower than both dehydrocrepenynate and the diacetylene.

While early tracer studies into most basidomycetes, and *F. hepatica* in particular, have given no cause to invoke biosynthetic steps other than the crepenynate pathway for compounds such as those in Figure 28, one report by Barley hints that the biochemistry of *F. hepatica* may be more complex [49]. Published incorporation studies on *F. hepatica* with positionally verified [9-^{14}C] labeled crepenynic acid found only 0.01% incorporation into the tetrayne, but found a 0.1% incorporation of labeled acetate [49]. Also, this study found a 200-fold increase in incorporation of labeled crepenynate into a partially characterized trienediol, relative to the tetrayne. From this data, the authors deduce that *F. hepatica* may possess an alternative, non-crepenynic acid pathway leading to polyaceteylene biosynthesis. Further incorporation studies with labeled 18:1$^{\Delta 9c,12a,14a}$ yielded even lower incorporation than crepenynic acid and may be indicative of diacetylenic precursor rearrangement or acyl chain degradation before subsequent desaturation and incorporation into polyacetylenes.

F. hepatica tissue is rich in oleate and linoleate, the precursors to crepenynate. Now that we have identified two enzymes that together are capable of producing $18{:}1^{\Delta 9c,12a,14a}$ from linoleate, there is strong evidence that the crepenynic acid pathway (Figure 28) does exist and is involved to some extent in polyacetylenic natural product biosynthesis. Our findings in no way discount or disprove the lack of crepenynate or diacetylene incorporation into the tetrayne seen by Barley. Rather, our findings further solidify Barley's hypothesis that a non-conventional crepenynic acid pathway may be involved in tetrayne production. By coupling Barley's results with those of the current experiment, you could make the argument that the biosynthesis of the ene-diyne follows the proposed crepenynic acid pathway, but may function as a shunt to an alternative pathway leading to certain polyacetylenic natural products. However, it is also conceivable that *Fistulina* might possess two different and distinct polyalkyne biogenesis pathways that act to manufacture different subsets of the polyacetylenes found therein; the first pathway is the established crepenynic acid pathway, as supported by this work, and a modified crepenynate derived pathway, which is supported by the 0.01% labeling of the tetrayne as described previously by Barley [49].

2.5 Conclusion

In conclusion, these experiments establish additional steps in the biosynthetic route for the formation of polyacetylenes in basidiomycetes. Herein, we described the functional characterization of two diverged desaturases isolated from *F. hepatica:* a Δ12-acetylenase and a novel trifunctional desaturase with Δ14-desaturase/acetylenase and Δ16-desaturase activities. These findings provide solid evidence that some *Fistulina* polyacetylenes are biosynthesized via the crepenynate pathway, a pathway proposed by Bu'Lock over half a century ago (Figure 28). Using the knowledge gained here, we have opened the door for the isolation of other diacetylenase gene homologs, which should provide more concrete evidence for the involvement of the crepenynic acid pathway in polyacetylene biosynthesis for other organisms, and may lead to the discovery of other

diverged desaturases capable of catalyzing later steps in the crepenynic acid pathway. Additionally, this enzyme provides another variation on the chemo- and regio-selectivity of the FAD2 enzyme family that will be interpretable when structural data ultimately becomes available.

CHAPTER 3. *ECHINACEA PURPUREA*

3.1 Background

Isolation and characterization of natural products found in the plant kingdom has uncovered a large diversity of unsaturated metabolites. However, the distribution of acetylene-containing metabolites has been shown to occur in only a limited number of plant families [33, 61, 62]. Some of the more notable polyacetylene producers belong to the families, Apiaceae, Araliaceae, Asteraceae, Santalaceae, and Umbelliferae. As of 2006, 32 plant species from Compositae, Leguminosae-Caesalpinioideae, and Rubiaceae had been found to contain crepenynic acid in their seed oil at up to 75% fatty acyl content [62]. This exotic fatty acid is known to be the precursor to a substantial number of polyacetylenes produced by plants and fungi [33, 34, 40, 49]. Crepenynic acid ($18{:}1^{\Delta 9c,12a}$) is the product of Δ12-acetylenase enzymes, which further desaturate the Δ12-double bond of linoleate. To date, most of the Δ12-acetylenase genes have been isolated from members of the families Apiaceae, Araliaceae, and Asteraceae.

Echinacea, commonly referred to as purple coneflower, is a perennial plant of the Asteraceae family. Echinacea are native to North America and have an extensive history as an herbal supplement and alternative medicine [63, 64]. For example, Native Americans commonly used *Echinacea* to treat a variety of ailments such as snakebites, fever, colds, and mouth sores. Today, despite lacking solid evidence for its efficacy, *Echinacea* is used as a dietary supplement to treat upper respiratory tract infections and the common cold. In previous studies, extracts have been reported to contain anti-inflammatory [65], anti-viral [66, 67], and anti-bacterial properties [65]. In addition to the immunomodulatory properites, recent data has emerged showing that *Echinacea*

extracts may contain anxiolytic properties [68, 69]. The most medicinally used species of *Echinacea* are *E. purpurea, E. angustifolia, and E. pallida*.

In many cases, polyacetylenes, alkylamides (alkamides) and other compounds have been implicated for the bioactivities of *Echinacea* [63, 67, 68, 70]. Alkamides make up a particularly interesting class of compounds that characteristically contain a fatty acid functionalized with an amine moiety, which is typically isobutylamine or 2-methylbutylamine in *Echinacea* (Figure 30). There have been at least 20 alkamides identified in *Echinacea* which contain varying degrees of unsaturation as exemplified by **1-5**. Alkamides can be found in a wide range of concentrations in different tissues of the plant. They have been measured at highest concentrations in the roots of *E. purpurea* and *E. angustifolia* and are less abundant in aerial parts [71].

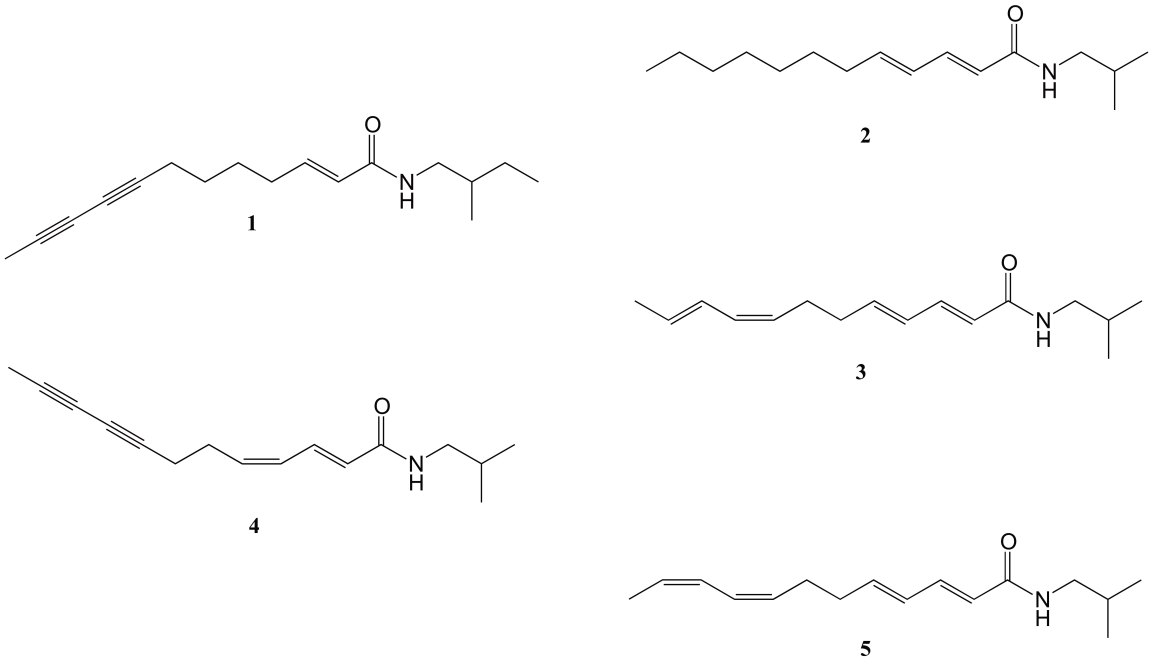


Figure 30: Examples of alkamides discovered in *Echinacea* [71].

Polyacetylenes in addition to acetylenic alkamides have been characterized from *Echinacea* (Figure 31). The bioactive role of these acetylenic metabolites is yet to be determined. Nevertheless, given the vast array of biological activities exhibited by other naturally occurring acetylenes (Chapter 1), it is believed that some of the unique

properties of *Echinacea* can be attributed to the presences of polyacetylenes. These polyacetylenes **6-10** have been isolated by several groups and are shown to contain as many as five carbon-carbon triple bonds (Figure 31). For a more in-depth look at the structural nature of polyacetylenes from *Echinacea* and other members of the Heliantheae tribe, readers are directed to the review by Christensen and Lam [72]. However, the goal of our study was not to investigate the mechanism or effects of *Echinacea* extracts but rather to try to elucidate the pathway from which the acetylenic natural products of the plant are derived.

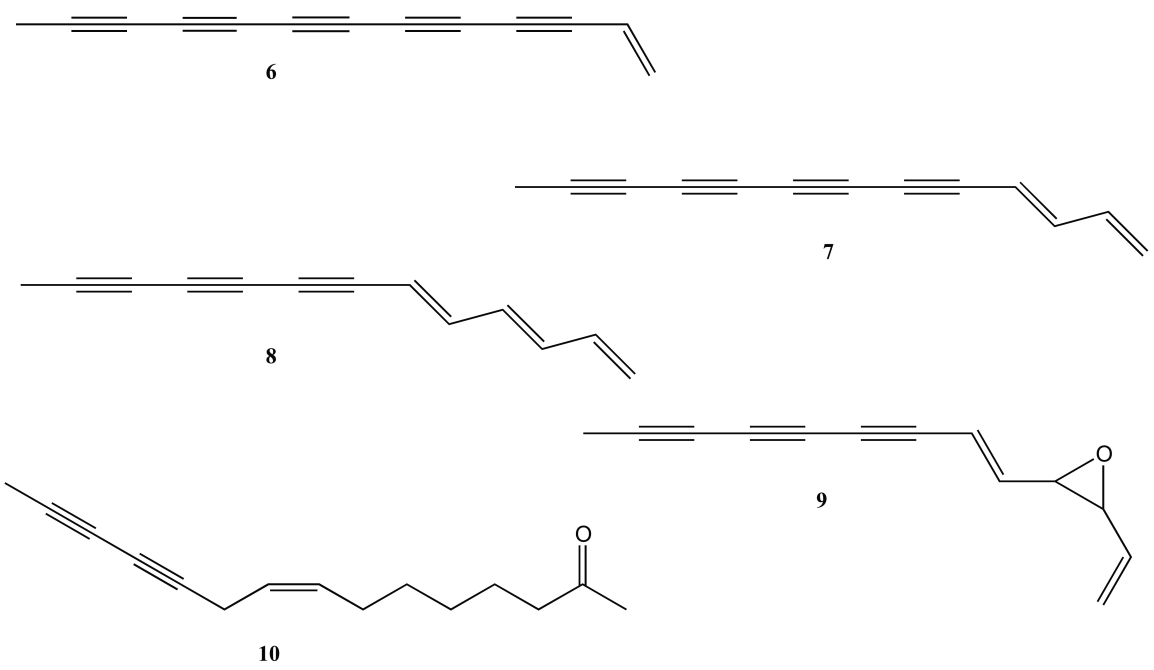


Figure 31: Examples of non-alkamide polyacetylenes found in *Echinacea* [63, 73, 74].

It is proposed that the polyacetylenes and acetylenic alkamides formed by *Echinacea*, as well as the vast majority of acetylenic specialized metabolites, originate from the crepenynic acid pathway. To better understand the process of acetylenic natural product synthesis in *E. purpurea,* this experiment was designed to test two distinct expression systems for the functional characterization of potential Δ12-acetylenases. The first of these systems, *Saccharomyces cerevisiae*, is one of the most

frequently employed eukaryotic expression systems. However, *S. cerevisiae* has some drawbacks when used for the expression of plant FAD2 diverged acetylenases. It is established that *Saccharomyces* lacks FAD2 activity and must be supplemented linoleate ($18:2^{\Delta 9c,12c}$) in order to see Δ12-acetylenase activity. A less trivial disadvantage of *Saccharomyces* is that, upon the expression of plant diverged epoxygenases and acetylenases, the system has provided little or no activity [25, 28]. The reason behind this lack of activity has remained elusive. For experiments investigating the multifunctional nature of the Δ12-acetylenases, namely the Δ12-desaturase activity acting on oleate, control through supplementation of polyunsaturated fatty acids in *S. cerevisiae* can be advantagous.

The second expression host is the oleaginous yeast, *Yarrowia lipolytica*. In recent years, *Y. lipolytica* has gained momentum as a potential source for renewable biofuels and a potent system for heterologous eukaryotic gene expression. This gaining momentum comes chiefly due to the "non-conventional" yeast's ability to accumulate polyunsaturated fatty acids and large lipid droplets. Experiments have shown both wild type and mutant strains capable of accumulating TAGs upwards of 30-50% of the cells dry weight [75]. An additional advantage of this expression system is that *Yarrowia* has a highly active FAD2 allowing for the accumulation of linoleate and enabling the direct characterization of putative Δ12-acetylenases to be accomplished without supplementation. Expression of the fungal Δ12-acetylenase CfACET indicated that *Y. lipolytica* can provide higher activity when expressing diverged FAD2 acetylenases (unpublished data, B. Blacklock).

In this project, we have uncovered the initial steps in the biosynthetic route for the formation of acetylenic natural products in *E. purpurea*. Using sequence homology to other functionally characterized plant Δ12-acetylenases, we targeted two putative acetylenases from *Echinacea* and, by heterologous expression in yeast, show their ability to catalyze the production of crepenynic acid, the first step in the formation of polyacetylenic natural products.

3.2 Methods and Materials

3.2.1 Isolation and Cloning of Potential Acetylenases from *Echinacea purpurea*

E. purpurea seedlings (accession # PI 631307; courtesy of North Central Regional Plant Introduction Station, Ames, Iowa) were grown on solid MS basal media with Gamborg's vitamins (0.05% MES, 1.45% sucrose, 0.22% MS basal medium, 0.2% v/v PPM, 1% agar, pH 5.75 with KOH) until they produced their first true leaves ("Stage 5 seedlings"), at which point, total RNA and mRNA isolation was carried out as described in Chapter 4. The full open reading frame of the two putative acetylenases, *Epa955* and *Epa2161*, was obtained through 3'- and 5'- RACE experiments using the SMARTer RACE cDNA Amplification and Advantage 2 PCR kits (Clontech, Mountain View, CA) with gene-specific primers (Table 15, Appendix B). These gene-specific primers were obtained from the original *E. purpurea* transcriptome data publically available on 03/21/2011. However, a second version of the transcriptome was made available to the public as of 10/07/2011 (http://medicinalplantgenomics.msu.edu/).

Once the complete open reading frames were identified, full-length transcripts containing a His_6 tag and 5'-*BamH*I and 3'-*Not*I restriction sites were obtained by *Taq* polymerase-mediated PCR (Table 15, Appendix B). PCR products were TOPO T/A cloned into pCR2.1 and transformed into DH5α chemically competent *E. coli*. The complete transcripts were subcloned into either pYES2 or pY5 vectors for expression in either *S. cerevisiae* or *Y. lipolytica*, respectively. Bidirectional sequencing was employed to ensure sequence integrity. Both silent and non-silent discrepancies between the PCR products and available transcriptomic sequences for *Epa955* and *Epa2161* were corrected using the Quik Change II site-directed mutagenesis kit (Stratagene, La Jolla, CA) and transformed into XL1-Blue chemically competent *E. coli* (Stratagene). Synthetic genes, codon-optimized for their respective expression systems, were purchased from Genscript (Piscataway, NJ). Sequences isolated through RACE experiments are

designated as EpaxxxRACE whereas those arising from the transcriptomal database are named EpaxxxTrans. The extenstion "-opt" is appened to codon-optimized sequences.

3.2.2 Heterologous Expression of *Echinacea purpurea* Constructs in Yeast

Constructs of *Epa955* and *Epa2161* were subsequently transformed into baker's yeast using a standard lithium acetate-polyethylene glycol method of Gietz [52] with selection on complete minimal medium lacking uracil (CM-ura+dex) for 3 days at 30 °C. The transformation of the yeast *Y. lipolytica* was carried out using a PEG/lithium acetate method with selection on CM-leu+dex for 5 days at 28 °C [76]. YPD start cultures were inoculated with single colonies and incubated at 28 °C for 24 h. Expression cultures for *S. cerevisiae* (CM-ura+2% gal) and *Y. lipolytica* (CM-leu+2% dex) were inoculated with 50 µL of starter culture and grown at 22 °C with 240 rpm for 24 h, followed by expression at 15 °C with 240 rpm for 24-72 h. For feeding experiments, exogenous fatty acids (provided as 95% ethanol stocks) and Tergitol NP-40 were supplemented to cultures at final concentrations of 0.5 mM and 0.1% (w/v), respectively.

3.2.3 GC/MS of Lipid Derivatives

FAME and DMOX derivatives from yeast cultures were prepared as previously described [54, 55]. In preparation for DMOX synthesis, the hexane solutions of FAMEs were initially evaporated to dryness and then reacted with 0.5 M NaOH in 90% aqueous MeOH at 80 °C for 1 h followed by acidification with 1 M HCl to generate the free fatty acids. The free fatty acids were washed with 1 mL of water and extracted upon addition of 2 mL of 1 : 1 hexane : diethyl ether. The organic layer containing the fatty acids was dried down under a stream of N_2 and the concentrated sample was used for DMOX derivatization.

The prepared FAMEs were separated and analyzed by an Agilent Technologies GC-MS (7890A GC/ 5975C MS) using a VF-23 column (30 m x 250 µm x 0.25 µm). For the analysis method "FAMEshort", the oven program was: 60 °C starting temperature, ramp

10 °C/min to 150 °C, hold 5 min; ramp 10 °C/min to 220 °C, hold 1 min; ramp 50 °C/min to 250 °C, hold 2 min. The "DMOX" oven program was the same program used to analyze FAMEs with the exception that the hold time for the 250 °C step was 12 min. In all cases the injector was set to 250 °C with a gas flow rates of 1.9 mL/min.

3.2.4 Bioinformatics

Transmembrane predictions were performed using the "TopPred 0.01" program from the Mobyle server (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::toppred). The multisequence alignments and phylogenetic analysis were performed as previously described by Minto [54]. Percent amino acid identity was determined using the sequence distance function of the MegAlign™ software from DNASTAR Inc. (Madison, WI).

3.3 Results

3.3.1 Isolation of Putative Δ12-Acetylenases

Using sequences of 48 known or hypothetical acetylenases and diverged desaturases from plants and fungi, a TBlastX search against the publically available *E. purpurea* transcriptome yielded two potential Δ12-acetylenase homologs, Epa955 and Epa2161. Although the transcriptome was available, RACE experiments were used to confirm the full extent of the ORF by the presence of a stop codon preceding the anticipated *fad2* start codon at the 5' end of the ORF and stop codon at the 3' end of the ORF followed by the poly $(A)^+$ tail for both diverged desaturases. Using the collected RACE data, full-length primers were designed with a His_6 tag and 5'-*BamH*I and 3'-*Not*I restriction sites. The use of the hexahistidine tag has been found to provide increased *in vivo* desaturation activity for the fusion protein in comparison to the native enzyme

[30, 54]. For each acetylenase construct, two His_6 fusion tags were used that contained His codons optimized for either the *S. cerevisiae* (CAT) or *Y. lipolytica* (CAC) expression systems.

Full-length targets were obtained by PCR amplification using *Taq* polymerase. Through comparison of the full-length sequence data with the transcriptome data, it was noticed that the PCR-derived sequences possessed likely SNPs (Table 6). These variations are expected because the PI 631307 accession is maintained as a genetically heterogenous population. SDM was used to correct the polymorphisms that result in translated sequence differences. The resultant constructs were sequenced and found to be in agreement with the RACE consensus. The complete ORF for each diverged desaturase was subcloned into both pYES2 and pY5 expression vectors and sequenced.

Table 6: Single-nucleotide polymorphs found in the putative Δ12-acetylenase constructs for *Echinacea purpurea*.

Sequence Sample	**Nucleotide #**	**Codon Change (RACE consensus→Observed)**	**Residue Change**
Epa 955	433	ACC → CCC	Thr → Pro
	458	GTT → GGT	Val → Gly
Epa 2161	68	GAT → GGT	Asp → Gly
	89	GAT → GGT	Asp → Gly

3.3.2 Bioinformatics

The targeted acetylenases *Epa955* and *Epa2161* contained 1140 and 1134 nt, respectively. The predicted 380-aa translation of Epa955 and the 378-aa sequence of Epa2161 were determined to share 84% identity to each other. The putative acetylenase, Epa2161 exhibited 93% sequence homology with the functionally

characterized Δ12-acetylenases from both *Helianthus annuus* and *Calendula officinalis*, whereas Epa955 has 83% sequence identity to the aforementioned acetylenases. A sequence alignment of Epa955 and Epa2161 with the functionally characterized FAD2s and diverged FAD2s established the presence of three highly conserved His motifs responsible for binding two active-site iron atoms required for catalysis (Figure 32). The conserved His box sequences for Epa955 and Epa2161 are: HECGH, HRNHH, and HVLHH. Bioinformatics experiments predicted the presence of four hydrophobic membrane-spanning regions for both putative acetylenases, which are characteristic of membrane-bound desaturases. Based on these data, it was anticipated that Epa955 and Epa2161 were functional Δ12-acetylenases.

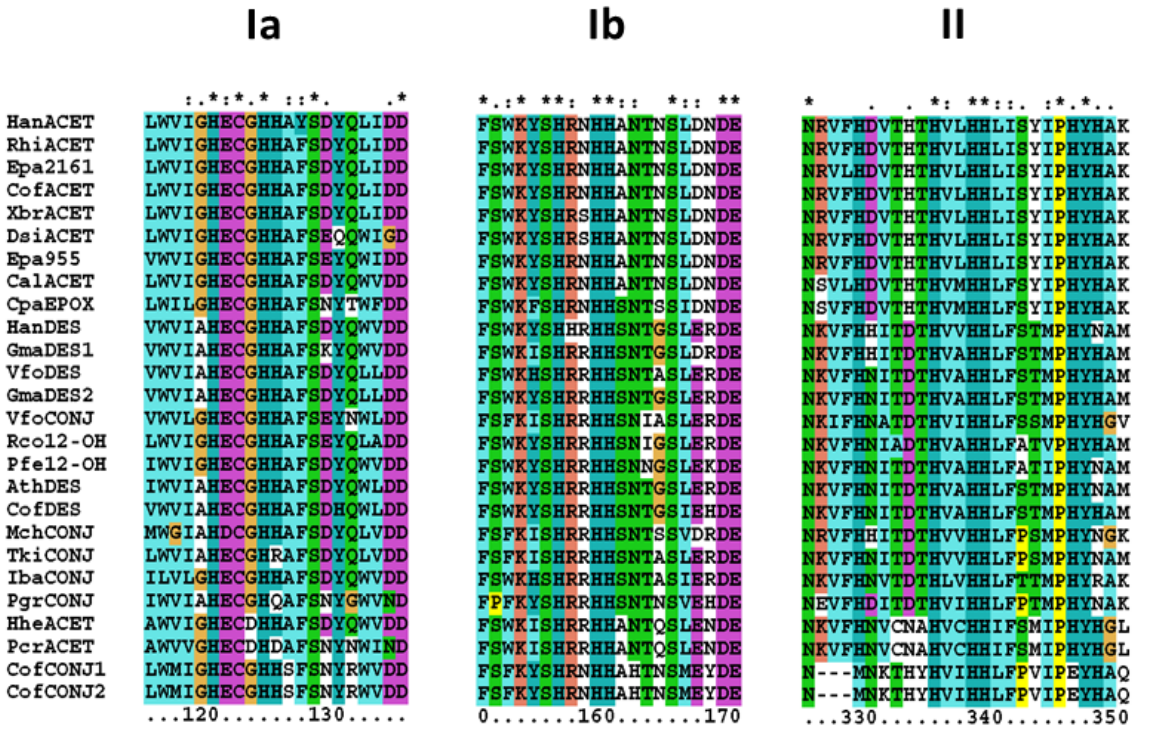


Figure 32: Local sequence alignment of the three conserved His boxes from diverged desaturases. The His box alignment contains sequences for FAD2 desaturases (DES), epoxygenases (EPOX), conjugases (CONJ), hydroxylases (12-OH), and acetylenases (ACET). The GenBank™ accession numbers for the displayed sequences are: CofACET, CAB64256.1; CofDES, AF343065; CofCONJ1, AF310155.1; CofCONJ2, AF310156; VfoDES, AF525534; VfoCONJ, AF525535; GmaDES1, P48630.1; GmaDES2, P48631; Rco12-OH, Q41131; Pfe12-OH, AAC32755; AthDES, P46313.1; HanDES, O24499; HanACET, AAO38032.1; MchCONJ, AF182521; CalACET, ABC00769.1; DsiACET, AAO38036.1; XbrACET, AAO38037.1, RhiACET, AAO38035.1; CpaEPOX, CAA76156.1; PgrCONJ, AAO37753.1; HheACET, AAO38031.1; TkiCONJ, AAO37751.1; PcrACET, AAB80697.1; and IbaCONJ, AAF05915. Sequence numbering corresponds to that of the multiple-sequence alignment.

3.3.3 Cloning and *Saccharomyces* Expression of Epa955 and Epa2161

The catalytic abilities possessed by the two *Echinacea* diverged desaturases were investigated through heterologous expression in the yeast, *S. cerevisiae*. For both Epa955 and Epa2161, the uncorrected RACE- and SDM-derived transcriptomic ORFs were subcloned into pYES2 and transformed into *S. cerevisiae* InvSc1 cells. For both the RACE -derived and transcriptomic sequences, the transgenic yeast expression cultures gave no detectable activity. To see if codon bias was playing a role in the lack of desaturase activity, a RACE-derived *Epa955* gene was synthesized in a codon-optimized form for *Saccharomyces*, transformed as a pYES2 construct, and expressed. Expression of the codon-optimized plant diverged desaturase again failed. These results extend the observations with previous expressions of unoptimized plant acetylenases and epoxygenases having little or no activity in *S. cerevisiae* expression systems [25, 28].

3.3.4 Cloning and *Yarrowia* Expression of Epa955 and Epa2161

Analogous to the experiments described with *S. cerevisiae*, RACE and transcriptomic sequences for both *Epa955* and *Epa2161* were expressed in *Y. lipolytica* behind the constitutive TEF promoter of the pY5 expression vector. Initial four-day expressions for all pY5 constructs revealed no activity in *Yarrowia* distinguishable from the empty vector constructs. A second, revised version of the *E. purpurea* transcriptome was subsequently published that had discrepancies with our original RACE-derived and transcriptomic sequences. The residue at 376 for original RACE-derived and transcriptome sequences was an Asp residue, while the second version of the transcriptome sequence had a Tyr residue. SDM was used to correct the Asp residue of the original constructs to Tyr376, but subsequent expressions of the corrected constructs still yielded no activity. An alignment of all of the expressed Epa955 gene translations highlighting the variations in amino acid sequences can be found in Figure 33.

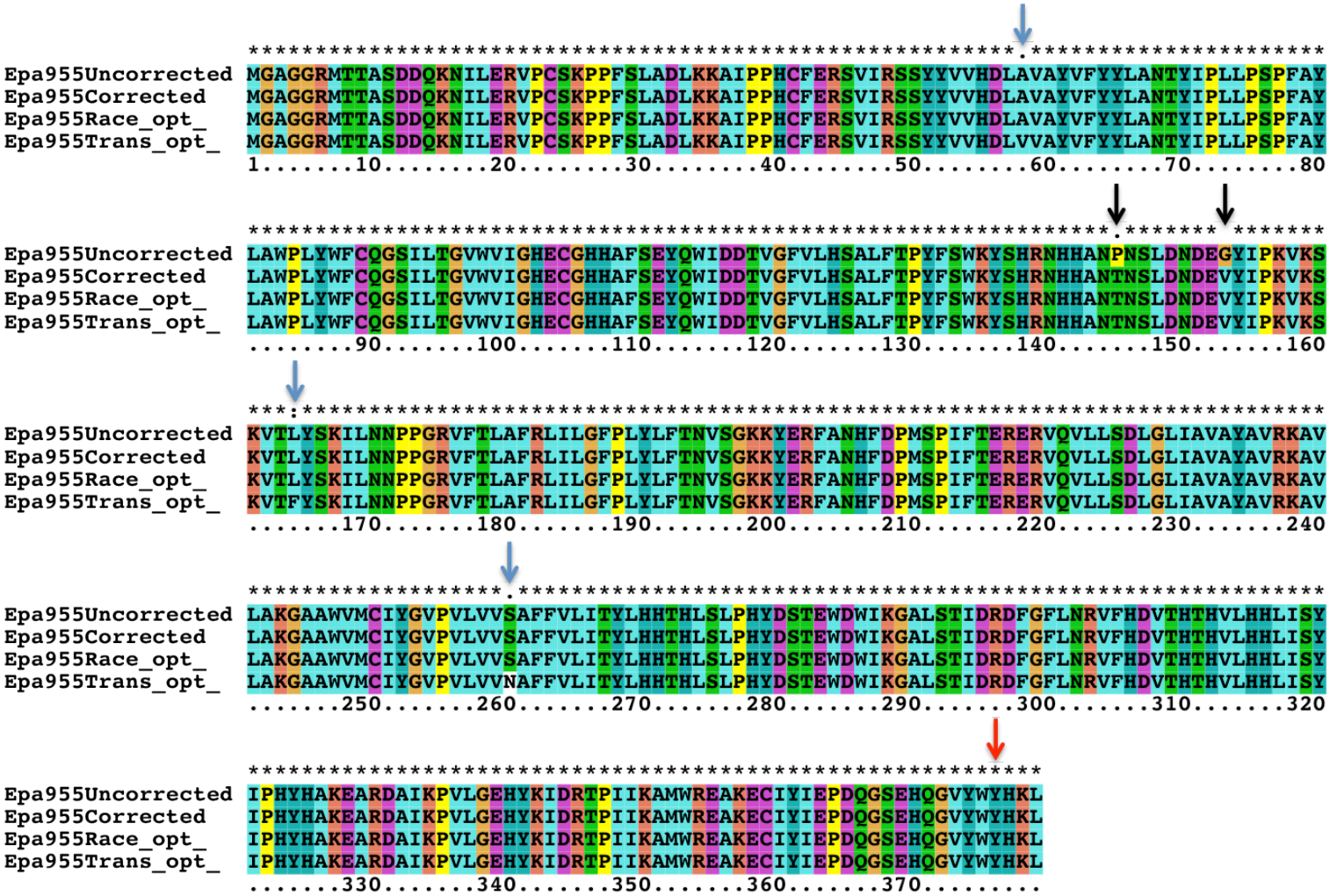


Figure 33: Multi-sequence alignment of all Epa955 expression constructs. The differences marked by black arrows represent the SNPs resulting in non-conservative primary sequence substitutions present in the original transcriptome sequence that were repaired by SDM to codons consistent with the RACE data and a broader alignment of FAD2-like sequences. Those marked by blue arrows represent the SNP changes between the Race-opt and Trans-opt constructs. The red arrow indicates the Try376 residue.

Unpublished work from our lab has indicated that *Yarrowia* may be more sensitive to codon bias than *Saccharomyces cerevisiae*. Only a synthetic, codon-optimized form of *CfACET* expressed in *Yarrowia* could be shown to have acetylenase activity. With this information in hand, synthetic, codon-optimized, transcriptomic constructs of each putative acetylenase were obtained and expressed in *Yarrowia.* GC-MS lipid analysis of the four-day, unfed *Y. lipolytica* expression cultures indicated the formation of three new fatty acid products by both Epa955Trans-opt and Epa2161Trans-opt (Figure 34). As the codon-optimized genes proved to be active, a second codon optimized gene was purchased for the RACE-derived Epa955 sequence, *Epa955RACE-opt*. The expression of this gene in *Y. lipolytica* also demonstrated the production of new fatty acids (Figure 34).

Further investigation uncovered that each expression experiment produced the same set of new fatty acids that were identifiable by their retention times and mass fragmentation patterns. The first of these fatty acids was found to accumulate to greater than 1.47% and was identified as $18{:}2^{\Delta 9c,12t}$, based on retention time (16.4 min) and mass fragmentation pattern. The mass spectra and retention time of the peak at 18.3 min proved to be identical to that of an authentic $18{:}1^{\Delta 9c,12a}$ methyl ester standard prepared from *Crepis* seed oil. The identity of the $18{:}1^{\Delta 9c,12a}$ formed by the transgenic yeast was further substantiated by comparison of DMOX derivatives of the analyte and a crepenynate standard. Interestingly, crepenynate was found to accumulate to essentially identical levels for both transcriptomic based constructs, ≈0.75% of the total lipid pool. However, the *Epa955RACE-opt* construct was found to accrue crepenynic acid to a slightly higher level (0.91%). It should be noted that the Epa955RACE-opt expression cultures also appeared to grow to a lower cell density over four days than the empty pY5 control, Epa955Trans-opt, and Epa2161Trans-opt expressions and this may have impacted the results of this experiment.

The third fatty acid product (retention time = 18.9 min) was found to be detectable only in trace amounts, making up ≤ 0.10 % of the yeast's total acyl content.

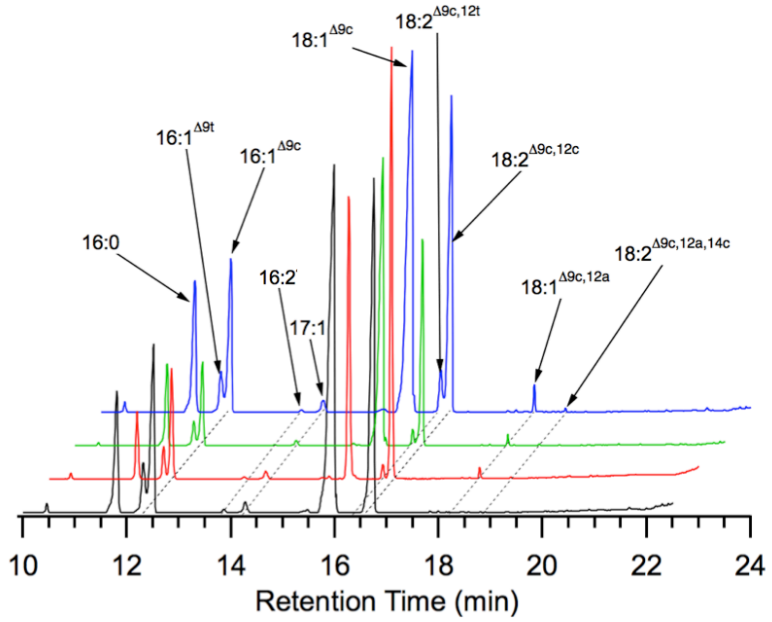


Figure 34: Overlay of four-day expressions of *Echinacea purpurea* genes in *Yarrowia.* The overlaid traces colored in black, red, green and blue correspond to GC-MS total ion chromatograms of total FAMES prepared from yeast expression cultures of empty pY5, pY5-$(His)_6$Epa955RACE-opt, pY5-$(His)_6$Epa955Trans-opt, and pY5-$(His)_6$Epa2161Trans-opt, respectively.

Table 7: Total fatty acid composition for four-day unfed expressions. Results are expressed as the mean % ± sd, n=4; *ND* = not detected. Growth conditions were 24 h at 22 ˚C; 240 rpm followed by 72 h at 15˚C; 240 rpm. [1] Expressed as pY5-$(His)_6$ constructs.

Fatty Acids →	**16:0**	**16:1**	**16:1**	**16:2**	**17:1**	**18:1$^{\Delta 9c}$**	**18:1$^{\Delta 11c}$**	**18:2$^{\Delta 9c,12t}$**	**18:2$^{\Delta 9c,12c}$**	**18:1$^{\Delta 9c,12a}$**	**18:2$^{\Delta 9c,12a,14c}$**
Constructs											
Unfed (4 day)											
pY5	9.88±0.41	4.12±0.98	12.3±0.9	0.12±0.09	0.74±0.22	43.6±0.8	0.18±0.23	*ND*	29.0±1.7	*ND*	*ND*
***Epa955Trans-opt*[1]**	10.2±0.4	3.0±0.27	10.5±0.4	*ND*	0.77±0.13	50.6±0.7	0.35±0.23	1.72±0.21	22.1±0.23	0.74±0.09	0.06±0.04
***Epa955RACE-opt*[1]**	7.76±0.33	3.97±0.20	12.8±0.4	0.33±0.02	1.32±0.09	32.6±2.0	0.42±0.05	1.47±0.03	38.0±1.6	0.91±0.05	0.10±0.01
***Epa2161Trans-opt*[1]**	10.2±0.3	3.31±0.30	10.8±0.9	0.05±0.06	0.96±0.15	47.9±2.4	0.10±0.19	2.89±0.26	23.0±1.2	0.75±0.08	0.06±0.05

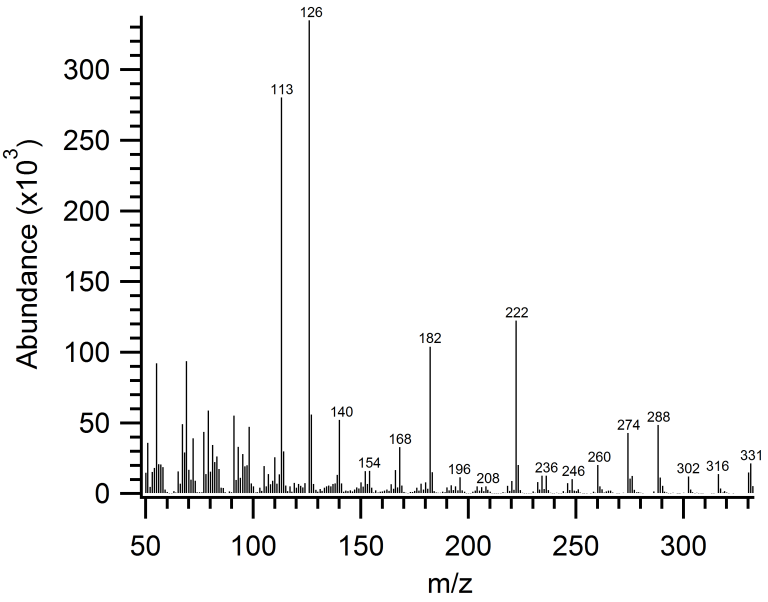


Figure 35: Mass spectra of crepenynate dimethyloxazoline prepared from Epa955Trans-opt expression cultures.

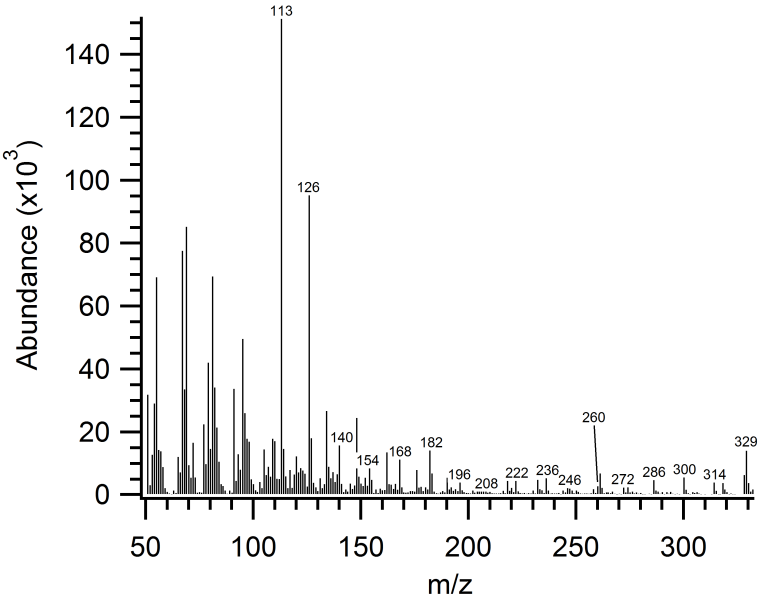


Figure 36: Mass spectra of dimethyloxazoline deriviative of dehydrocrepenynic acid prepared from Epa955Trans-opt expression cultures.

This second analyte was confirmed to be dehydrocrepenynic acid (18:2$^{\Delta 9c,12a,14c}$) through comparison of relative retention times and mass spectra with authentic FAMEs and DMOX standards made from *C. cibarius* fruit bodies. The origin of this acetylenic metabolite may stem from two potential sources. The formation of the conjugated Δ14-double bond could arise from endogenous desaturase activity possessed by Epa955 and Epa2161, and would therefore be a new example of multifunctional catalytic activity found in diverged desaturases. However, a more probable explanation for the biosynthesis of this fatty acid is the endogenous and highly active FAD2 of *Y. lipolytica* carrying out the desaturation at carbon 14, consistent with the work of Blacklock [30].

3.3.5 Feeding Experiments in *Yarrowia lipolytica*

In an effort to elucidate the origin of dehydrocrepenynic acid, attempts were made to supplement crepenynate to the *Yarrowia* expression cultures. However, this task proved to be a daunting one. The addition of 18:1$^{\Delta 9c,12a}$ at 0.5 mM was surprisingly toxic to *Yarrowia*, given that *Saccharomyces* liquid cultures are routinely supplemented at the same concentration. In an attempt to identify the maximum amount of 18:1$^{\Delta 9c,12a}$ that could be supplemented, the effect of exogenous crepenynate on liquid cultures grown at 28 °C were observed. Acceptable cell growth was found with exogenous 18:1$^{\Delta 9c,12a}$ at or below 0.1 mM (Figure 37). Although the conditions for the toxicity curve were different than that of the expression conditions, it was believed that the toxicity curve would provide valuable insight into potential feeding conditions.

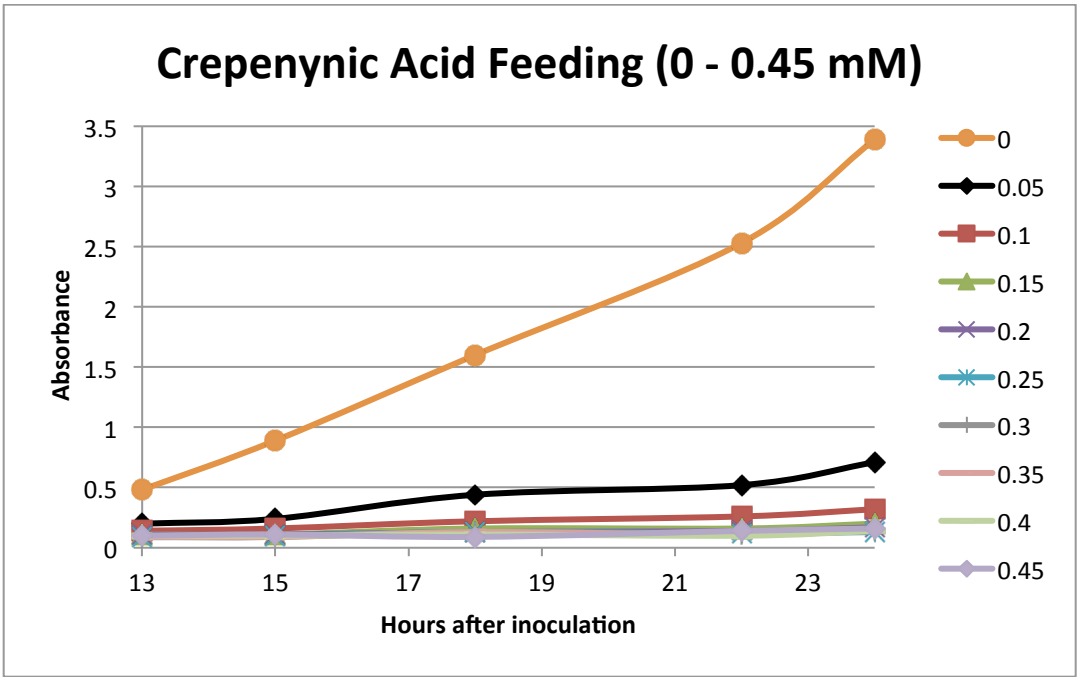


Figure 37: Toxicity curve for crepenynic acid supplementation. The yeast cultures (CM-leu+2% dex+ 0.1% (w/v) NP-40) were fed exogenous fatty acid and grown at 28 °C with 240 rpm shaking for 24 h.

In order to circumvent the cumulative effects of reduced culture growth and the low temperature incubations typically required for high FAD2 activities, 0.1-0.5 mM crepenynic acid was supplemented to the empty pY5 control expression cultures after the initial growth period at 22 °C for 24 h, followed by 24 h at 15 °C, as opposed to the original conditions of 15 °C for 72 h. It was discovered that $18{:}1^{\Delta 9c,12a}$ accumulated to reasonable levels at all supplemented concentrations, with the best being seen at 0.5 mM ($18{:}1^{\Delta 9c,12a}$ and $18{:}2^{\Delta 9c,12a,14c}$ at ≈ 24% and ≈ 0.83% of total fatty acids, respectively). Based on the observed results, $18{:}1^{\Delta 9c,12a}$ feedings were setup and grown at 22°C for 24 h. At this point, crepenynate was supplemented to a final concentration of 0.5 mM and gene expression was carried out at 15 °C for 24 h. Additionally, to insure that the unfed and supplemented expressions were carried out under the same conditions, unfed two-day expressions were performed in conjunction.

The unfed expression cultures were noticeably less dense after two days relative to the four-day expressions. Furthermore, all of the expression cultures appeared to have grown to nearly equal optical densities. The transgenic lipid pools were found to

contain >0.69% 18:2$^{\Delta 9c,12t}$ for the RACE construct and 2-4.5 times that for the transcriptomic constructs (Table 8). FAMEs analysis exhibited a substantially higher accumulation of both crepenynate (2-fold) and dehydrocrepenynate (6-fold) for Epa955Trans-opt and Epa2161Trans-opt, in comparison to the four-day expressions (Table 8). Surprisingly the Epa955RACE-opt sample demonstrated half the accumulation of crepenynate and over a 50% increase of dehydrocrepenynic acid than in the four-day expression. These apparently inconsistent data remain an unresolved question.

FAMEs analysis of the two-day 18:1$^{\Delta 9c,12a}$ feeding expressions demonstrated a decrease in 18:2$^{\Delta 9c,12t}$ formation for Epa955Trans-opt and Epa2161Trans-opt, but Epa955RACE-opt had a slight increase relative to the two-day unfed expressions. The build-up of crepenynic and dehydrocrepenynic acids was detected in the lipid pool of all supplemented cultures (Figure 38; Table 8). Crepenynate was found to accumulate >19% of the total FAMEs, but with substantial variability, while dehydrocrepenynate accrued to >0.2%, but more consistently. In all supplemented acetylenase expressions, the amount of 18:1$^{\Delta 9c,12a}$ and 18:2$^{\Delta 9c,12a,14c}$ stockpiles were 7% less than and 0.1% greater than the empty pY5 control, respectively.

FAMEs prepared from six *E. purpurea* seeds were analyzed by GC-MS to check for the accumulation of any acetylenic compounds (Figure 39). Although the seed lipid pool contained a significant amount of linoleate, the known Δ12-acetylenase substrate, no acetylenic natural products were identified.

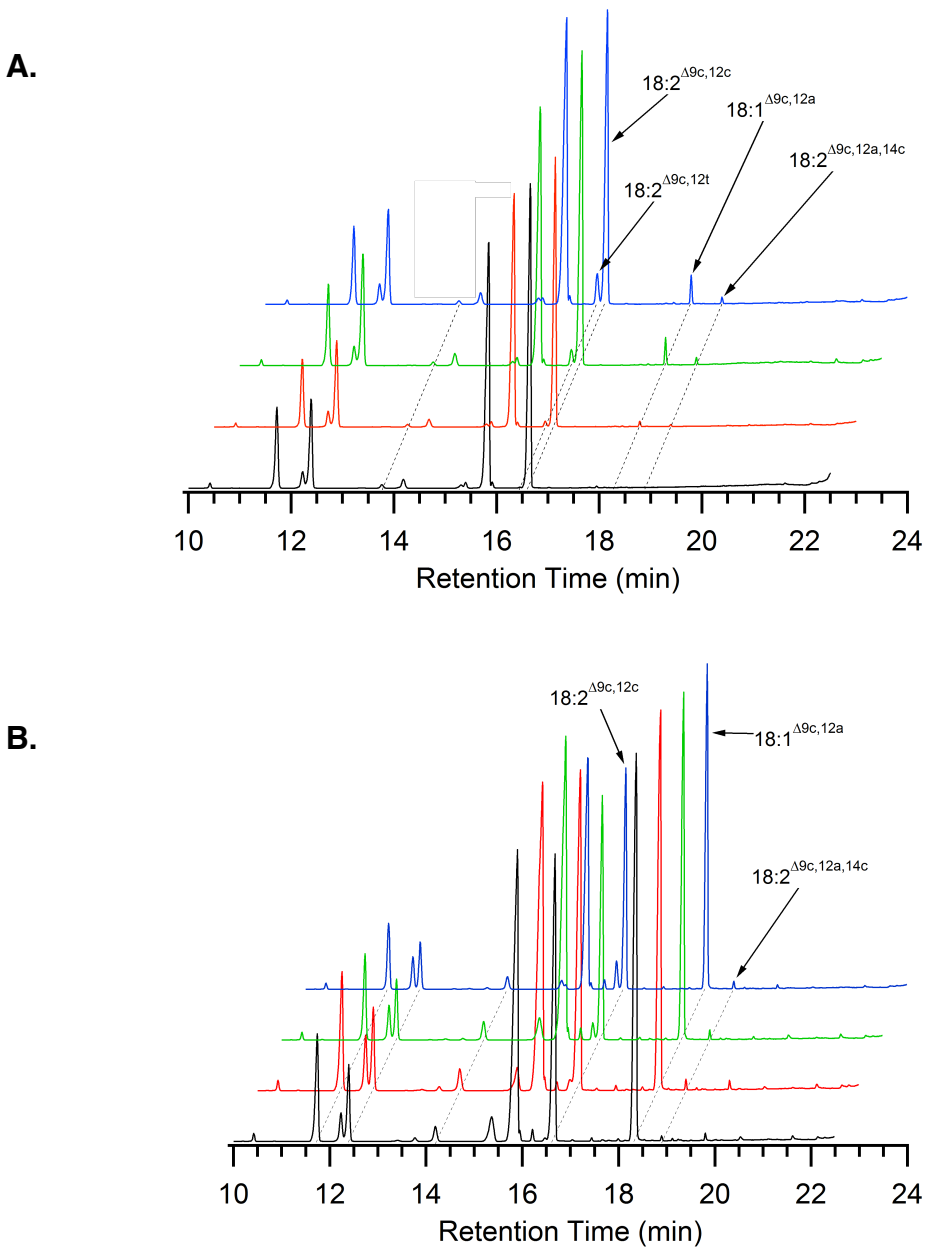


Figure 38: Overlay of two-day expressions of *Echinacea purpurea* Δ12-acetylenases in *Yarrowia*. (A) represents the unfed expressions while (B) corresponds to the samples supplemented with 0.5-mM crepenynic acid. The overlaid traces colored in black, red, green, and blue correspond to total FAMES analyzed by GC-MS for yeast expression cultures of empty pY5, pY5-$(His)_6$Epa955RACE-opt, pY5-$(His)_6$Epa955Trans-opt, and pY5-$(His)_6$Epa2161Trans-opt, respectively.

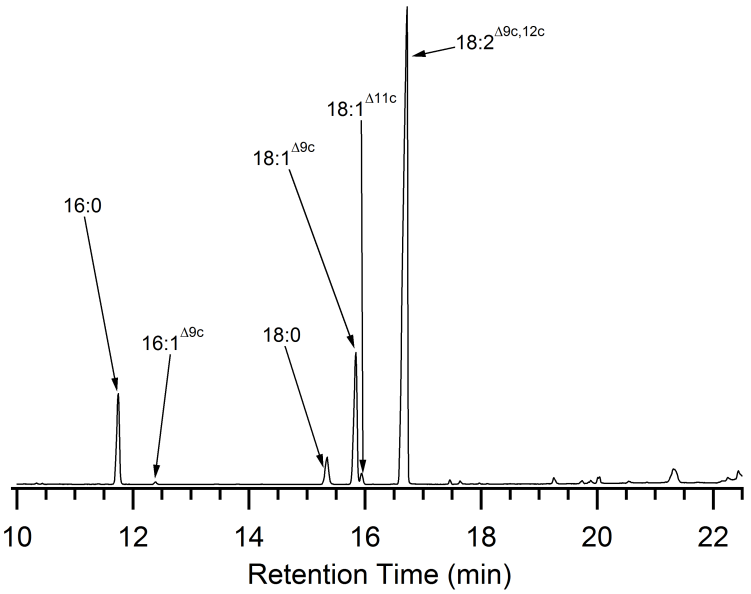


Figure 39: GC-MS analysis of FAMEs isolated from *Echinacea purpurea* seeds. No acetylenic fatty acid accumulation was observed.

Table 8: Tabulated data for unfed and crepenynic acid supplemented two-day *Echinacea* acetylenase-expressing cultures. Results are expressed as the mean % ± sd, n=4; *ND* = not detected. Cultures were grown with 240 rpm shaking at 22 ˚C for 24 h followed by 15 ˚C for 24 h. [(a)]Fatty acid supplemented to expression culture after the initial 24 h growth at 22 ˚C. [1] Expressed as pY5-$(His)_6$ constructs.

Fatty Acids →	**16:0**	**16:1**	**16:1**	**16:2**	**17:1**	**18:1$^{\Delta 9c}$**	**18:1$^{\Delta 11c}$**	**18:2$^{\Delta 9c,12t}$**	**18:2$^{\Delta 9c,12c}$**	**18:1$^{\Delta 9c,12a}$**	**18:2$^{\Delta 9c,12a,14c}$**
Constructs											
Unfed (2 day)											
pY5	9.66±0.37	2.21±0.29	11.4±1.8	0.47±0.14	1.31±0.14	39.1±6.0	0.49±0.05	0.09±0.01	34.6±4.4	*ND*	*ND*
Epa955*Trans-opt*[1]	9.07±0.45	2.67±0.29	12.2±1.0	0.39±0.07	1.48±0.23	38.2±2.8	0.46±0.04	1.69±0.15	31.5±3.4	1.36±0.27	0.35±0.13
Epa955*RACE-opt*[1]	9.12±0.31	2.27±0.25	12.2±1.2	0.39±0.06	1.29±0.04	38.2±2.3	0.46±0.03	0.69±0.03	34.2±1.1	0.39±0.07	0.16±0.02
Epa2161*Trans-opt*[1]	8.60±0.54	2.07±0.46	10.5±1.8	0.39±0.05	1.62±0.08	38.7±1.8	0.55±0.03	3.09±0.62	31.5±0.5	1.50±0.18	0.38±0.03
+18:1$^{\Delta 9c,12a}$ *(a)*											
pY5	6.84±0.67	2.66±1.30	4.80±0.29	0.25±0.04	1.25±0.13	28.2±3.3	0.36±0.03	0.20±0.03	23.8±2.0	29.8±4.2	0.21±0.01
Epa955*Trans-opt*[1]	6.71±0.32	4.11±2.14	4.27±0.79	0.15±0.03	1.46±0.25	34.2±4.1	0.43±0.03	1.37±0.20	21.6±3.8	22.9±8.1	0.43±0.11
Epa955*RACE-opt*[1]	7.33±0.10	5.08±1.71	3.81±0.81	0.20±0.05	1.62±0.08	34.3±3.9	0.30±0.09	0.71±0.04	25.0±0.8	19.1±5.1	0.30±0.03
Epa2161*Trans-opt*[1]	6.85±0.69	5.22±1.61	4.57±0.58	0.17±0.03	1.72±0.43	33.2±3.7	0.57±0.11	2.38±0.27	21.9±3.8	21.6±5.1	0.40±0.08

3.4 Discussion

In this chapter, we described the cloning and functional characterization of two diverged desaturases isolated from *Echinacea purpurea* seedlings, and investigated an alternative yeast expression system for the characterization of plant diverged desaturases.

Plants and higher fungi are some of the most prolific producers of acetylenic metabolites. These bioactive metabolites have been isolated from a limited number of plant families, and the pathway leading to their biosynthesis is not well understood. However, the vast majority of acetylenic secondary metabolites are believed to originate from the branchpoint compound crepenynic acid, the product of a FAD2-like Δ12-acetylenase that introduces an alkynyl group at C12 of linoleate [40]. To date, most of the Δ12-acetylenases have been isolated from Apiaceae, Araliaceae, and Asteraceae. As with other Asteraceae species, flowers in the genus *Echinacea* have been found to produce a number of acetylene-containing compounds. In addition to acetylenic alkamides, *Echinacea* manufactures tri-, tetra-, and pentaacetylenic compounds that are purported to derive from crepenynate.

Our investigation began with a TblastX search through the publically available *E. purpurea* transcriptome and produced a list of putative Δ12-acetylenases. The top two candidates, Epa955 and Epa2161, contained the characteristic hydrophobic domains of membrane-desaturases and the trademark His motifs involved in binding a catalytically active non-heme di-iron complex within the active site. Likewise, the two diverged desaturases were believed to be functional acetylenases, as they possessed substantial sequence identity (>80%) with other functionally characterized Δ12-acetylenases from *H. annuus* and *C. officinalis.*

Two distinct yeast expression systems were used in an effort to functionally characterize Epa955 and Epa2161. Upon expression of the anticipated acetylenases in the first system, *S. cerevisiae*, no activity was detected. This result is consistent with previous attempts of expressing plant acetylenases in *Saccharomyces* [25, 28]. Although the reason for this lack of activity is unclear, this work rules out question of codon bias.

Initial expression of both acetylenase constructs in *Yarrowia* also proved to be unsuccessful. Because the expression of a codon-optimized version of the *CfACET* gene for *Y. lipolytica* had higher activity than a non-optimized transcript, it was surmised that a critical parameter in *Yarrowia* may be codon bias. Cultures expressing codon-optimized versions of both Epa955 and Epa2161 were discovered to accumulate $18{:}1^{\Delta 9c,12a}$ and trace amounts of $18{:}2^{\Delta 9c,12a,14c}$. Although functional characterization of plant Δ12-acetylenases has been performed with transgenic plant seeds, there are many advantages of using *Y. lipolytica,* such as the relative ease of transformation, low media expense, short cultivation time, and high TAG accumulation. Coupled together, the results obtained herein suggest that *Yarrowia* provides a suitable alternative expression system for plant Δ12-acetylenases and that codon bias needs to be considered during construct design.

In contrast to the expression of Epa955Trans-opt, it is important to note that initial expressions of the *Epa955RACE-opt* construct indicated no desaturation or acetylenation activity. It was realized that the sequence for the *Epa955Trans-opt* construct was obtained from the second version of the published *E. purpurea* transcriptome, while the sequence for the *Epa955RACE-opt* gene was obtained from RACE data and the original transcriptome. Analysis of the sequence data indicated the presence of a non-silent mutation at residue 376. It was discovered that the RACE and original transcriptome contained Asp376 while the second version of the transcriptome possessed Tyr376. By conducting a BlastX search with the Epa955 sequence, the conservation of Tyr376 was confirmed by the top 100 hits. Furthermore, this region of the gene has the characteristics of a typical ER targeting sequence (–YXXL–) [77, 78]. SDM was carried out to convert both optimized and unoptimized *Epa955* constructs to Tyr376. Expression of the resultant *Epa955RACE-opt* construct was the only construct to provide Δ12-acetylenase activity.

The production of crepenynate and dehydrocrepenynate demonstrates that both target genes contain Δ12-acetylenase catalytic abilities. Likewise, another example of FAD2-like enzymatic multifunctionality is illustrated here. Previously characterized

homologs have displayed Δ12-desaturase activity resulting in the production of 16:2$^{\Delta 9c,12t}$, 18:2$^{\Delta 9c,12t}$, and 18:2$^{\Delta 9c,12c}$ [19, 30]. In this case, however, the diverged desaturases were observed to specifically convert oleate into 18:2$^{\Delta 9c,12t}$. Although the 18:2$^{\Delta 9c,12c}$ fatty acid is present in the transgenic yeast lipid pool, the accumulation of this product in Epa955 and Epa2161 is equal to or below the levels measured for the empty pY5 control. The exception to this is the four-day expression of Epa955RACE-opt, where the level of 18:2$^{\Delta 9c,12c}$ was found to accumulate to a substantially higher amount than the control. Although, it must be stated that the Epa955RACE-opt expression cultures grew to a lower cell density than the other expression cultures.

To test for additional activities that would be consistent with the crepenynate pathway, our functionally characterized *E. purpurea* Δ12-acetylenases were supplemented 0.5 mM-crepenynic acid. Although, *S. cerevisiae* are routinely supplemented at the same concentration, 0.5 mM-crepenynate was surprisingly toxic to *Yarrowia.* Information gathered from a toxicity curve indicated that *Y. lipolytica* had limited growth in the presence of 0.05-0.1 mM crepenynate, and supported the need for pre-growing cultures before crepenynic acid supplementation. By initially growing the cultures for 24 h at 22 °C and then supplementing exogenous crepenynate, it was discovered that after 24 h at 15 °C the yeast cells had taken up, and even utilize the exogenous crepenynate to produce dehydrocrepenynate.

The finding that 18:2$^{\Delta 9c,12a,14c}$ is produced in the empty pY5 *Y. lipolytica* cells when they are provided crepenynic acid is consistent with the observations of Blacklock [30]. This further solidifies the catalytic role of FAD2 in the formation of 18:2$^{\Delta 9c,12a,14c}$. In contrast to the similar accumulation levels of 16:2$^{\Delta 9c,12t}$ and 18:2$^{\Delta 9c,12c}$ described above, there is a 50%-100% increase in the build-up of dehydrocrepenynate found in the supplemented two-day acetylenase expression cultures in comparison to the control. Given that the level of 18:1$^{\Delta 9c,12a}$ incorporation was lower in the acetylenase-expressing cultures than in the control, and the formation of dehydrocrepenynic acid is higher in these cultures, it appears that Epa955 and Epa2161 also possess weak Δ14-desaturase activity. An alternative explanation for the increased level of 18:2$^{\Delta 9c,12a,14c}$ in the

acetylenase expression cultures could be due to metabolic channeling of *de novo* $18{:}1^{\Delta 9c,12a}$ to the endogenous FAD2 of *Yarrowia,* resulting in higher dehydrocrepenynic acid accumulation. However to confirm this activity, Epa955 and Epa2161 would need to be expressed in a mutant strain of *Yarrowia* with FAD2 knocked out.

3.5 Conclusion

We have described the isolation and characterization of the first two bifunctional Δ12-acetylenases from *Echinacea.* These two diverged desaturases contain the catalytic machinery to carry out Δ12-desaturation/acetylenation. Although further studies are required to provide a direct link between these Δ12-acetylenases and the formation of polyacetylenes and acetylenic alkamides, our results provide enzymatic evidence supporting *E. purpurea* acetylenic metabolite production occurs via the crepenynic acid pathway (Figure 40). Additionally, we have shown that *Y. lipolytica* is a suitable expression system for plant diverged desaturases, and provide yet another case for the value of *Yarrowia* as a heterologous expression system.

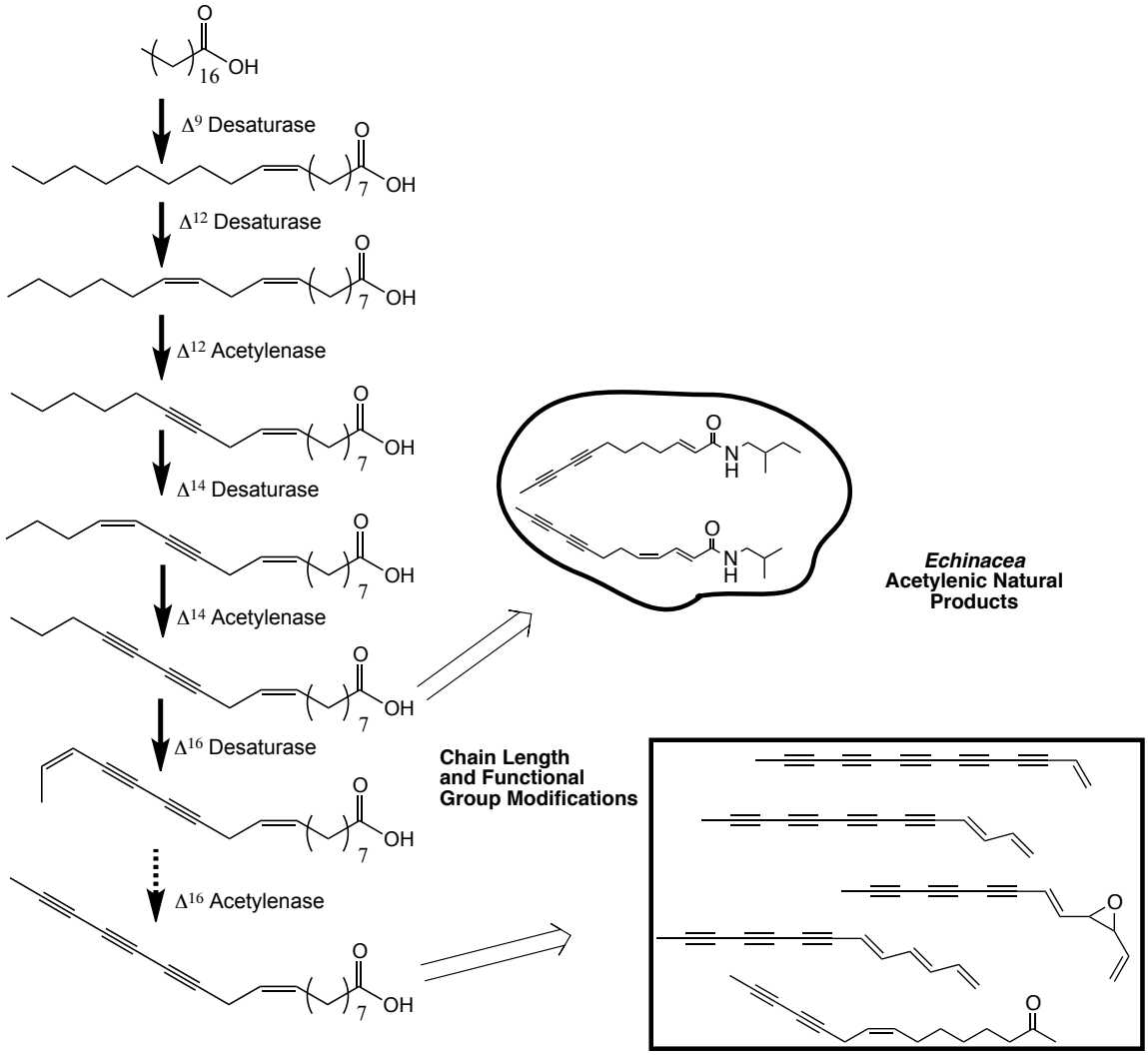


Figure 40: Crepenynic acid pathway with proposed branch points for the formation of polyacetylenic natural products in *Echinacea.*

CHAPTER 4. MATERIALS AND METHODS

4.1 Materials

All yeast and *E. coli* media components were purchased from Becton Dickinson & Co. (Franklin Lakes, NJ), unless stated otherwise. Tergitol NP-40 and galactose were acquired from Sigma-Aldrich (St. Louis, MO). NaCl and NaOH were ordered from Fisher Scientific (Fair Lawn, NJ). All restriction enzymes and buffers were purchased from New England Biolabs (Ipswich, MA). All RNA work was performed using RNase-free reagents and equipment. Water used for RNA protocols was treated with 0.1% v/v diethylpyrocarbonate (DEPC) by stirring overnight at ambient temperature and subsequently autoclaving for 20 min at 121 °C to deactivate DEPC. All primers were obtained from Integrated DNA Technologies (Coralville, IA) and synthetic genes were purchased from GenScript (Piscataway, NJ). *Fistulina hepatica* fruit bodies were a gracious gift from Professor D. Job (University of Neuchâtel, Switzerland). *Echinacea purpurea* seedlings were provided by the North Central Regional Plant Introduction Station, Ames, Iowa.

4.2 Methods

4.2.1 Isolation of Gene Targets

4.2.1.1 Total RNA Isolation from Tissue

By a method similar to the suggested protocol for using the TRIzol (Invitrogen, Carlsbad, CA) reagent, total RNA was isolated from *Fistulina hepatica* mycelium and

fruiting bodies (ATCC strain 64428), as well as *E. purpurea* seedlings (accession # PI 631307). Fresh tissue was collected and ground to a fine powder in mortar with a pestle cooled by a liquid nitrogen bath. Once ground, 1 mL of TRIzol per 50-100 mg of tissue was added to the mortar and the tissue was allowed to thaw in the presence of the reagent. One microliter of the mixture was aliquoted into 2-mL screw cap vials (containing ≈ 500 µL of 0.5-mm glass beads) and disrupted with a Bio spec products mini-beadbeater-8 at max speed for 90 s. The solution was allowed to cool on ice for 90 s, agitated vigorously as above, and left at room temperature for 5 min. Centrifugation was performed at 12,000×*g* for 10 min at 4 °C, and the supernatant was transferred to a clean 1.7-mL tube. Chloroform (200 µL) was added, the mixture was vortexed for 30 s, and left at ambient temperature for 3 min. The solution was spun as above for 15 min and the aqueous layer was transferred to a new 1.7-mL tube. Next, 250 µL of isopropanol and 250 µL of citrate solution (0.8 M Na citrate & 1.2 M NaCl) were added to the aqueous layer that was inverted and centrifuged as above for 15 min. The supernatant was removed and the pellet was washed with 75% EtOH. The pellet was allowed to air-dry at 37 °C for 5-10 min and then dissolved in 20-30 µL of 10 mM Tris-HCl (pH 7.5).

4.2.1.2 Isolation of Messenger RNA from Total RNA

Poly $(A)^+$ RNA was isolated from total RNA using the Dynabeads mRNA Purification Kit (Invitrogen). The total RNA solution (32-75 µg in a minimum of 100 µL volume) was heated to 65 °C for 2 min, and immediately transferred to ice for 3-5 min. The Dynabeads were resuspended in the storage buffer by vortexing and 200 µL of the slurry was transferred to a microfuge tube. The beads were extracted from the suspension by a magnet while the supernatant was removed by a pipet. After the tube was removed from the magnet, the beads were washed with 100 µL of Binding Buffer (20 mM Tris-HCl, pH 7.5; 1.0 M LiCl; 2 mM EDTA) with vortexing. Using the magnet assembly, the wash supernatant was removed and a volume of Binding Buffer equal to

the starting solution volume of total RNA was added to the tube. The beads were resuspended, and the total RNA solution was added, vortexed, and placed on a roller/mixer for 5 min at room temperature. The Dynabeads containing the bound mRNA were extracted from the solution with the magnet. The beads were washed twice with 200 µL of Washing Buffer B (10 mM Tris-HCl, pH 7.5; 0.15 M LiCl; 1 mM EDTA). Finally, the Dynabeads were resuspended in 20 µL of 10-mM Tris-HCl (pH 7.5), heated to 72 °C for 2 min, and immediately placed on the magnet. The supernatant containing isolated mRNA was recovered and stored at -80 °C.

4.2.1.3 Synthesis of cDNA from mRNA

Amplification of Target cDNA from mRNA--- Using mRNA obtained from the Dynabead mRNA Purification Kit, reverse transcription was performed using the Superscript First-Strand Synthesis system for RT-PCR (Invitrogen). The reverse transcription reactions were set up in 1.7-mL microfuge tubes according to Table 9.

Table 9: Setup for reverse transcription reactions.

Component	**+RT**
Isolated mRNA	6.0-8.0 µL
10-mM dNTP mix	1.0 µL
$Oligo(dT)_{12-18}$ (0.5 µg/µL)	1.0 µL
DEPC-treated H_2O	to a final volume of 10 µL

Each reaction mixture was incubated at 65 °C for 5 min and cooled on ice for an additional 3-5 min. A master solution containing 2 µL of 10x RT buffer (200 mM Tris-HCl, pH 8.4; 500 mM KCl), 4 µL of 25 mM $MgCl_2$, 2 µL of 0.1 M DTT, and 1 µL of RNaseOUT (40 U/µL) per reaction was prepared. For each reaction, 9 µL of the master solution was added, mixed, and incubated at 42 °C for 2 min. Then, 1 µL of SuperScript II RT was added to the sample and the mixture was incubated on a heating block at 42 °C for ≈50 min. The temperature was increased to 70 °C for 15 min and the reaction was chilled on

ice for 2-3 min. After spinning briefly, 1 µL of RNase H was added to the sample, which was then incubated at 37 °C for 20 min.

4.2.1.4 Generation of RACE-Ready cDNAs

Using the SMARTer RACE cDNA Amplification Kit (Clontech, Mountain View, CA), rapid amplification of cDNA ends (RACE) for both 5' and 3' ends was performed. A master mix was prepared containing 5 µL of 5x First-Strand buffer (250 mM Tris-HCl, pH 8.3; 375 mM KCl; 30 mM $MgCl_2$), 2.5 µL of 20 mM DTT, and 2.5 µL of dNTP mix (10 mM each). The synthesis of 5' and 3' first-strand products were carried out in separate tubes. The 5' sample contained 2.75 µL of mRNA and 1 µL of 5'-CDS Primer A, where as the 3' sample had 3.75 µL of mRNA and 1 µL of 3'-CDS Primer A. The two tubes were mixed, incubated at 72 °C for 3 min, and then cooled to 42 °C for 2 min. The tubes were centrifuged at 14,000×*g* for 10 s to collect the reaction mixture, and 1 µL of SMARTer IIA oligo was added to the 5' reaction only. Next, 0.625 µL of 40 U/µL RNase inhibitor and 2.5 µL of 100 U/µL SMARTScribe reverse transcriptase was added to the master mix, mixed, and 5.25 µL of reverse transcriptase solution was added to both the 5' and 3'-RACE-Ready cDNA tubes. After mixing by pipetting and a brief centrifugation step to collect the samples, they were incubated at 42 °C for 90 min followed by a 70 °C heating step for 10 min. Finally, the samples were diluted with 20-100 µL of Tricine-EDTA buffer (10 mM Tricine-KOH, pH 8.5; 1 mM EDTA) depending on the starting mRNA concentration.

4.2.1.5 Rapid Amplification of cDNA Ends

Rapid amplification of cDNA ends was performed using the RACE-Ready cDNA products as a template. Each of the 5' and 3' RACE samples were setup according to Table 10. The master mix was prepared containing 34.5 µL of PCR-grade water, 5.0 µL of 10x Advantage 2 PCR buffer (400 mM Tricine-KOH, pH 8.7 at 25 °C; 150 mM KOAc; 35 mM $Mg(OAc)_2$; 37.5 µg/mL BSA; 0.05% Tween 20; 0.05 % Nonidet-P40), 1.0 µL of dNTP

mix (10 mM each), and 1.0 µL of 50x Advantage 2 polymerase mix (50% glycerol; 15 mM Tris-HCl, pH 8.0; 75mM KCl; 0.05 mM EDTA) per reaction. Genomic DNA (75-100 ng) was used for control reaction.

Table 10: RACE amplification tables. The top table represents the set up for 5' RACE amplification, while the bottom table shows the components of the 3' RACE amplification. GSP corresponds to the gene specific primer used for amplification while UPM is the universal primer mix. The components for each amplification reaction were added in the order shown.

Component	5' Race Sample (µL)	GSP1/GSP2 (µL)	UPM only (µL)	GSP1 only (µL)	gDNA control (µL)
5' RACE cDNA	2.5	2.5	2.5	2.5	2.5 of gDNA
UPM (10x)	5	--	5	--	--
5'-RACE Primer (10 µM)	1	1	--	1	1
3'-RACE Primer (10 µM)	--	1	--	--	1
PCR-Grade H_2O	--	4	1	5	4
Master Mix	41.5	41.5	41.5	41.5	41.5

Component	3' Race Sample (µL)	GSP1/GSP2 (µL)	UPM only (µL)	GSP2 only (µL)	gDNA control (µL)
3' RACE cDNA	2.5	2.5	2.5	2.5	2.5 of gDNA
UPM (10x)	5	--	5	--	--
5'-RACE Primer (10 µM)	--	1	--	--	1
3'-RACE Primer (10 µM)	1	1	--	1	1
PCR-Grade H_2O	--	4	1	5	4
Master Mix	41.5	41.5	41.5	41.5	41.5

RACE amplification was performed on an Eppendorf Mastercycler Gradient 5331 Thermal Cycler using the following program:

RACE

1. 94 °C, 30 s
2. 72 °C, 3 min
3. Go to 1; Repeat 5 times
4. 94 °C, 30 s
5. 70 °C, 30 s
6. 72 °C, 3 min
7. Go to 4; Repeat 5 times
8. 94 °C, 30 s
9. 68 °C, 30 s
10. 72 °C, 3 min
11. Go to 8; Repeat 25 times

4.2.1.6 *Taq* Polymerase Chain Reaction Amplification

Using the reverse transcription products, PCR reactions using the GoTaq Flexi polymerase kit (Promega, Madison, WI) were developed to amplify the target genes for *Echinacea*. Each of the PCR samples contained the components found within the table below. PCR amplification was performed using an Eppendorf Gradient Cycler using a temperature program designated "REMTD".

Table 11: GoTaq polymerase meditated amplification of gene targets.

Component	**Volume (µL)**
5x Green GoTaq Buffer	4
25 mM $MgCl_2$	1.6
10 mM dNTP	0.4 of each dNTP
GoTaq Polymerase (5U/µL)	0.1
10 µM primer	1.0
DNA template	1-2 (depending on concentration)
H_2O	Volume up to 20

REMTD

1. 94 °C, 2 min
2. 94 °C, 15 s
3. 72 °C, 15 s; Increment step temperature by -0.5 °C
4. 72 °C, 3.5 min
5. Go to 2; Repeat 28 times
6. 94 °C, 15 s
7. 58 °C, 15 s
8. 72 °C, 3.5 min
9. Go to 6; Repeat 10 times
10. 72 °C, 7.5 min

4.2.1.7 Polymerase Chain Reaction Amplification of *Fistulina hepatica* Gene Targets by Vent Polymerase

Using the reverse transcription products obtained for *F. hepatica*, PCR reactions were developed to amplify the target genes. Each of the PCR samples contained 2 µL of 10x IDT Buffer (pH 8.3; 20 mM $MgCl_2$), 0.4 µL of each 10 mM dNTP solution, 0.2 µL Vent polymerase, 1.0 µL of template, 1.0 µL of each of the respective primers (10 µM), and 13.2 µL of H_2O. PCR amplification was performed using an Eppendorf Gradient Cycler and the program "REMTD".

4.2.1.8 Gel Purification of PCR Products

Using the QIAEX II Gel extraction kit protocol by QIAgen (Valencia, CA), the DNA product was purified from the agarose gel. DNA bands were extracted from the gel with a razor blade and Buffer QXI was added in a ratio of 300 µL per 0.1 g of gel slice. For DNA greater than 4 kb in size, an additional 200 µL of water was added per 0.1 g. The QIAEX II matrix was resuspended and 10 µL was added to each DNA isolation mixture. The sample was incubated at 50 °C for 10 min with vortexing every 2 min to ensure that the DNA and agarose were suspended in solution. After the 10 min incubation period, the mixture was centrifuged at $16{,}100 \times g$ (maximum speed) for 30 s and the

supernatant was removed. The remaining pellet was washed once with 500 µL of Buffer QXI and twice with 500 µL of Buffer PE by resuspension and centrifugation. The pellet was allowed to air-dry at 37 °C for 5-10 min, followed by resuspension of the pellet in 20 µL of TE (10 mM Tris-HCl, pH 8.0; 1 mM EDTA) by vortexing. For DNA of less than 4 kb, this elution period was at room temperature for 5 min. For DNA greater than 4 kb, elution was carried out at 50 °C for 5 min. After the incubation, the sample was spun at maximum speed for 30 s and the supernatant containing the DNA of interest was transferred to a clean tube.

4.2.2 Cloning and Transformation of *Escherichia coli*

4.2.2.1 Topoisomerase-Mediated Cloning and Transformation into *Escherichia coli*

The genes of interest were cloned into the vector pCR2.1 using a TOPO TA Expression kit (Invitrogen, Carlsbad, CA). Half-size cloning reactions relative to the manufacturer's instructions were setup according to Table 12. If Vent polymerase was used for full-length amplification, the overhangs required for TA cloning were generated by incubating the Vent polymerase-mediated PCR products with 0.25 µL of 5 U/ µL GoTaq polymerase at 72 °C for 10 min.

Table 12: General setup for TOPO T/A cloning reactions.

Components	**Component volume (µL)**	**Half Reactions (µL)**
Gene product	0.5 to 4.0	2.0
1.2 M NaCl, 0.06 M $MgCl_2$	1.0	0.5
pCR 2.1 TOPO TA Vector	1.0	0.5
MB H_2O	Volume up to 6.0	-

After flicking to mix and collecting by brief centrifugation, the cloning reactions were incubated at room temperature for 15-25 min. Vials of DH5α chemically competent *E. coli* cells were thawed on ice and the entire cloning reaction was added to a thawed vial of cells and flicked to mix. The vials were incubated on ice for 5-30 min.

After incubating, the vials were heat-shocked for 30 s in a water bath equilibrated to 42 °C, and then chilled on ice. Room temperature SOC medium (250 µL) was added to each sample followed by incubation at 37 °C for 1 h with 200 rpm shaking. Two volumes (25 µL and 50 µL) of the transformation reaction was plated on separate Luria-Bertani broth plates (LB) + ampicillin (amp) and grown overnight at 37 °C. Each of the LB plates (1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar) contained 100 µg/mL ampicillin that allowed for the selection of positive transformants based on antibiotic screening. Single colonies were picked from each plate and inoculated into 5-mL liquid cultures of LB + antibiotic. These liquid cultures were incubated with shaking (250 rpm) at 37 °C for 16-18 h.

4.2.2.2 Subcloning into Yeast Expression Vectors

The pCR 2.1 constructs and circularized expression vectors, pY5 and pYES2, were digested with *BamH*I and *Not*I restriction enzymes. The digests were incubated at 37 °C for approximately 2 h, at which point the linearized expression vectors were treated with Antarctic phosphatase at 37 °C for an additional 1 h. The reactions were incubated at 65 °C for 15 min to heat-inactive the restriction and phosphatase enzymes. The digest samples were ran on an agarose gel and the linearized products were then purified by the QIAEX II Gel extraction kit.

The ligation of gene inserts into yeast expression vectors was carried out using T4 DNA ligase (Gibco BRL). Each ligation reaction contained 4:1 molar equivalence of insert : vector, 4 µL of 5x ligase buffer, 2 µL of 10 mM ATP, 1 µL of ligase, and molecular-biology grade H_2O to a final volume of 20 µL. The reactions were ligated for 1-3 h at room temperature and 10 µL of each ligation reaction was transformed into XL1-Blue *E. coli* and screened on LB + amp plates.

4.2.2.3 Transformation of XL1-Blue *Escherichia coli*

Transforming DNA (0.5-10 µL; dependent on starting concentration) was added to 100 µL of freshly thawed RbCl chemically competent XL1-Blue *E.coli* and incubated on ice for 15-30 min [53]. The cell solution was heat-shocked in a water bath at 42 °C for 45 s and then placed back on ice for an additional 3 min. LB media (1 mL) was added to the transformation reaction and the culture was incubated at 37 °C for 1 h with occasional inversion to mix. The cells were then pelleted at 3,000*xg* for 1 min before all but 150 µL of the supernatant was discarded. The pelleted cells were gently resuspended in the remaining supernatant and spread on LB + amp plates.

4.2.3 Site-Directed Mutagenesis

Each reaction using the Quik Change II site-directed mutagenesis kit (Stratagene) contained the components below, and the components were added in the order shown. SDM was performed using the MutLong program below. Once the thermocycler program had finished, the template DNA was digested with 0.5 µL of *Dpn*I (10 U/µL) at 37 °C for 1 h. An aliquot of the digested product (2 µL) was then used to transform RbCl competent, XL1-Blue *E. coli*.

Table 13: Site-directed mutagenesis reactions.

Components	**Component amount**
10x Reaction buffer (100 mM KCl, 100 mM $(NH_4)_2SO_4$, 200 mM Tris-HCl, 20 mM $MgSO_4$, 1% Triton X-100, 1 mg/mL BSA)	2.5 µL
plasmid DNA	25 ng
SDM primer 1	62.5 ng
SDM primer 2	62.5 ng
dNTP mix	0.5 µL
MB H_2O	Volume up 25 µL
Pfu Ultra High-Fidelity DNA polymerase (2.5 U/uL)	0.5 µL

MutLONG

1. 95 °C, 30s
2. 95 °C, 30s
3. 55 °C, 1 min
4. 68 °C, 9.5 min
5. Go to 2; Repeat 16 times

4.2.4 Yeast Expressions of DNA Constructs

4.2.4.1 *Saccharomyces cerevisiae* Expression

Gene products for *S. cerevisiae* expression were cloned into the expression vector pYES2. These constructs were subsequently transformed into the *S. cerevisiae* strain INVSc1 using a standard lithium acetate-polyethylene glycol method of Gietz [52]. The transformed *S. cerevisiae* was grown on complete minimal medium lacking the nucleobase, uracil, containing dextrose (CM-ura+dex) for 3 days at 30 °C [53]. YPD liquid starter cultures (1% Bacto Yeast Extract, 2% Bacto peptone, and 2% dextrose) from single transformant colonies were grown at 30 °C overnight with 240 rpm shaking. Expression cultures for the INVSc1 transformants were generated by inoculating 50 µL of the YPD start cultures into 5-mL of CM-ura+2% galactose. The expression cultures were grown at 22 °C with 240 rpm for 24 h followed by expression at 15 °C with 240 rpm for 72 h. In the event that exogenous fatty acids were fed, the 5-mL expression culture contained 0.1% w/v Tergitol NP-40 and the fed fatty acid at a final concentration of 0.5 mM. All fatty acid feeding stocks were prepared as 10 mg/mL solutions in 95% ethanol.

4.2.4.2 *Yarrowia lipolytica* Expression

In preparation for *Yarrowia* expression, the full-length target transcripts were subcloned into the pY5, a *Y. lipolytica* expression vector containing the *Leu2* marker. The transformation of the *Y. lipolytica* Leu2 auxotroph (ATCC 76982) was carried out

using a PEG/lithium acetate method [76]. The transgenic yeast was cultivated on CM-leu+2% dextrose for 3-5 days at 28 °C, and single transformant colonies were used to innoculate YPD start cultures. Expression of the gene constructs was carried out as before in CM-leu+2% dex and were grown at 22 °C with 240 rpm for 24 h, followed by expression at 15 °C with 240 rpm for 24-72h. For feeding experiments, exogenous fatty acids and Tergitol NP-40 were supplemented to final concentrations of 0.5 mM and 0.1% (w/v), respectively.

4.2.4.3 Fatty Acid Derivatization of Transgenic Yeast Expressions

4.2.4.3.1 Fatty Acid Methyl Esters (FAMEs)

FAMEs from both transformed yeast species were prepared by incubation of resuspended cell pellets in 2% H_2SO_4 in methanol (1 mL) at 80 °C for 1-2h. If the transformants were fed exogenous fatty acids, the cells were first washed with 1% (w/v) Tergitol NP-40 (2 x 5 mL), followed by two washes of H_2O (5 mL) before the addition of acidic methanol. The FAMEs were extracted into hexanes (2 mL) and concentrated under a stream of N_2. The concentrated FAMEs were then redissolved in 200-500 µL of hexanes.

4.2.4.3.2 Fatty Acid Pyrrolidides (FAPys)

Fatty acid pyrrolidides were developed by a reacting 250 µL of pyrrolidine and 25 µL of glacial acetic acid with a dried FAMEs sample at 100 °C for 1 h. The pyrrolidides were recovered by adding 1 mL of water and 2 mL of hexane : diethyl ether (1 : 1). The products were then washed twice with 2 mL of H_2O, and the organic phase was dried down under a nitrogen atmosphere to concentrate the fatty acid pyrrolidide derivatives. The pyrrolidides were analyzed by GC-MS after being redissolved in 200 µL of hexanes.

4.2.4.3.3 Fatty Acid Dimethyloxazolines (DMOX)

FAMEs samples were concentrated under a stream of N_2 and then reacted with 0.5 M NaOH in MeOH : H_2O (90 : 10) at 80 °C for 1 h followed by acidification with 1 M HCl. As derivatizations of higher unsaturated fatty acids were performed, this step was carried out with the 0.5 M NaOH solution at room temperature, overnight. The free fatty acids were extracted with 2 mL of hexanes and concentrated under N_2.

The concentrated free fatty acids were reacted with 750 µL of distilled DCM : DMF solution (1000 : 1) and 500 µL of oxayl chloride under nitrogen to generate the acid chloride. Once the reaction stopped bubbling, the solvent was dried down and the acid chlorides were reacted with 500 µL of 10 mg/mL 2-amino-2-methylpropanol in distilled DCM. The mixture was allowed to react for 1 h at room temperature under a nitrogen atmosphere. The solution was dried down and reacted with 1000 µL of trifluoroacetic acid at 40 °C for 1 h to generate fatty acid DMOX derivatives. The derivatized fatty acids were dried down and recovered by addition of 1 mL of H_2O and extraction with hexane (2 x 2 mL). The organic layer was then washed twice with H_2O, and concentrated under a stream of nitrogen. Subsequently, the DMOX derivatives were dissolved in 50-200 µL of hexanes and analyzed by GC-MS

REFERENCES

REFERENCES

1. Rustan, A.C. and C.A. Drevon, *Fatty Acids: Structures and Properties.* Encycl Life Sci, 2005. 1-7
2. Kliewer, S.A., S.S. Sundseth, S.A. Jones, P.J. Brown, G.B. Wisely, C.S. Koble, P. Devchand, W. Wahli, T.M. Willson, J.M. Lenhard, and J.M. Lehmann, *Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors α and γ.* Proc Natl Acad Sci U S A, 1997. 94:4318-23.
3. Ohlrogge, J. and J. Browse, *Lipid biosynthesis.* Plant Cell, 1995. 7:957-70.
4. Okuley, J., J. Lightner, K. Feldmann, N. Yadav, E. Lark, and J. Browse, *Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis.* Plant Cell, 1994. 6:147-58.
5. Pereira, S.L., A.E. Leonard, and P. Mukerji, *Recent advances in the study of fatty acid desaturases from animals and lower eukaryotes.* Prostaglandins Leukot Essent Fatty Acids, 2003. 68:97-106.
6. Fox, B.G., J. Shanklin, C. Somerville, and E. Munck, *Stearoyl-acyl carrier protein Δ-9 desaturase from Ricinus communis is a diiron-oxo protein.* Proc Natl Acad Sci U S A, 1993. 90:2486-90.
7. Shanklin, J., J.E. Guy, G. Mishra, and Y. Lindqvist, *Desaturases: emerging models for understanding functional diversification of diiron-containing enzymes.* J Biol Chem, 2009. 284:18559-63.
8. Shanklin, J. and E.B. Cahoon, *Desaturation and related modifications of fatty acids.* Annu Rev Plant Physiol Plant Mol Biol, 1998. 49:611-41.

9. Sperling, P., P. Ternes, T.K. Zank, and E. Heinz, *The evolution of desaturases.* Prostaglandins Leukot Essent Fatty Acids, 2003. 68:73-95.

10. Shanklin, J., E. Whittle, and B.G. Fox, *Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase.* Biochemistry, 1994. 33:12787-94.

11. Minto, R.E., P.R. Adhikari, and G.A. Lorigan, *A 2H solid-state NMR spectroscopic investigation of biomimetic bicelles containing cholesterol and polyunsaturated phosphatidylcholine.* Chem Phys Lipids, 2004. 132:55-64.

12. Minto, R.E., W.J. Gibbons Jr, T.B. Cardon, and G.A. Lorigan, *Synthesis and conformational studies of a transmembrane domain from a diverged microsomal Δ12-desaturase.* Anal Biochem, 2002. 308:134-40. 13. Coon, M.J., *Omega oxygenases: nonheme-iron enzymes and P450 cytochromes.* Biochem Biophys Res Commun, 2005. 338:378-85.

14. Shanklin, J., C. Achim, H. Schmidt, B.G. Fox, and E. Münck, *Mössbauer studies of alkane ω-hydroxylase: evidence for a diiron cluster in an integral-membrane enzyme.* Proc Natl Acad Sci U S A, 1997. 94:2981-6.

15. Shanklin, J. and E. Whittle, *Evidence linking the Pseudomonas oleovorans alkane ω-hydroxylase, an integral membrane diiron enzyme, and the fatty acid desaturase family.* FEBS Lett, 2003. 545:188-92.

16. Jacobson, B.S., J.G. Jaworski, and P.K. Stumpf, *Fat metabolism in higher plants: LXII. stearyl-acyl carrier protein desaturase from spinach chloroplasts.* Plant Physiol, 1974. 54:484-6.

17. Buist, P.H. and B. Behrouzian, *Deciphering the cryptoregiochemistry of oleate Δ12 desaturase: a kinetic isotope effect study.* J Am Chem Soc, 1998. 120:871-6.

18. Behrouzian, B., L. Fauconnot, F. Daligault, C. Nugier-Chauvin, H. Patin, and P.H. Buist, *Mechanism of fatty acid desaturation in the green alga Chlorella vulgaris.* Eur J Biochem, 2001. 268:3545-9.

19. Carlsson, A.S., S. Thomaeus, M. Hamberg, and S. Stymne, *Properties of two multifunctional plant fatty acid acetylenase/desaturase enzymes.* Eur J Biochem, 2004. 271:2991-7.

20. Evangelista, R.L., *Oil extraction from Lesquerella seeds by dry extrusion and expelling.* Ind Crops Prod, 2009. 29:189-96.

21. Jaworski, J. and E.B. Cahoon, *Industrial oils from transgenic plants.* Curr Opin Plant Biol, 2003. 9:178-84.

22. Dyer, J.M., D.C. Chapital, J.C. Kuan, R.T. Mullen, C. Turner, T.A. McKeon, and A.B. Pepperman, *Molecular analysis of a bifunctional fatty acid conjugase/desaturase from tung. Implications for the evolution of plant fatty acid diversity.* Plant Physiol, 2002. 130:2027-38.

23. van de Loo, F.J., P. Broun, S. Turner, and C. Somerville, *An oleate 12-hydroxylase from Ricinus communis L. is a fatty acyl desaturase homolog.* Proc Natl Acad Sci U S A, 1995. 92:6743-7.

24. Broun, P., S. Boddupalli, and C. Somerville, *A bifunctional oleate 12-hydroxylase: desaturase from Lesquerella fendleri.* Plant J, 1998. 13:201-10.

25. Lee, M., M. Lenman, A. Banas, M. Bafor, S. Singh, M. Schweizer, R. Nilsson, C. Liljenberg, A. Dahlqvist, P.O. Gummeson, S. Sjodahl, A. Green, and S. Stymne, *Identification of non-heme diiron proteins that catalyze triple bond and epoxy group formation.* Science, 1998. 280:915-8.

26. Reed, D.W., D.R. Polichuk, P.H. Buist, S.J. Ambrose, R.J. Sasata, C.K. Savile, A.R. Ross, and P.S. Covello, *Mechanistic study of an improbable reaction: alkene dehydrogenation by the Δ12 acetylenase of Crepis alpina.* J Am Chem Soc, 2003. 125:10635-40.

27. Kirsch, C., K. Hahlbrock, and I.E. Somssich, *Rapid and transient induction of a parsley microsomal Δ-12 fatty acid desaturase mRNA by fungal elicitor.* Plant Physiol, 1997. 115:283-9.

28. Cahoon, E.B., J.A. Schnurr, E.A. Huffman, and R.E. Minto, *Fungal responsive fatty acid acetylenases occur widely in evolutionarily distant plant families.* Plant J, 2003. 34:671-83.

29. Fritsche, K., E. Hornung, N. Peitzsch, A. Renz, and I. Feussner, *Isolation and characterization of a calendic acid producing (8,11)-linoleoyl desaturase.* FEBS Lett, 1999. 462:249-53.

30. Blacklock, B.J., B.E. Scheffler, M.R. Shepard, N. Jayasuriya, and R.E. Minto, *Functional diversity in fungal fatty acid synthesis: the first acetylenase from the Pacific golden chanterelle, Cantharellus formosus.* J Biol Chem, 2010. 285:28442-9.

31. Sperling, P., M. Lee, T. Girke, U. Zahringer, S. Stymne, and E. Heinz, *A bifunctional Δ^6-fatty acyl acetylenase/desaturase from the moss Ceratodon purpureus. A new member of the cytochrome b5 superfamily.* Eur J Biochem, 2000. 267:3801-11.

32. Serra, M., B. Pina, J.L. Abad, F. Camps, and G. Fabrias, *A multifunctional desaturase involved in the biosynthesis of the processionary moth sex pheromone.* Proc Natl Acad Sci U S A, 2007. 104:16444-9.

33. Bohlmann, F., T. Burkhardt, and C. Zdero, *Naturally Occurring Acetylenes*, 1973. London: Academic Press.

34. Minto, R.E. and B.J. Blacklock, *Biosynthesis and function of polyacetylenes and allied natural products.* Prog Lipid Res, 2008. 47:233-306.

35. Bu'Lock, J.D., D.C. Allport, and W.B. Turner, *The biosynthesis of polyacetylenes. Part III. Polyacetylenes and triterpenes in Polyporus anthracophilus.* J Chem Soc, 1961. 1654-7.

36. Allen, E.H. and C.A. Thomas, *Trans-trans-3,11-tridecadiene-5,7,9-triyne-1,2-diol, an antifungal polyacetylene from diseased safflower (Carthamus tinctorius).* Phytochemistry, 1971. 10:1579-82.

37. Guillet, G., B.J.R. Philogène, J. O'Meara, T. Durst, and J.T. Arnason, *Multiple modes of insecticidal action of three classes of polyacetylene derivatives from Rudbeckia hirta.* Phytochemistry, 1997. 46:495-8.

38. Fernandez-Lopez, R., C. Machon, C.M. Longshaw, S. Martin, S. Molin, E.L. Zechner, M. Espinosa, E. Lanka, and F. de la Cruz, *Unsaturated fatty acids are inhibitors of bacterial conjugation.* Microbiology, 2005. 151:3517-26.

39. Christensen, L.P., W. Vach, J. Ritskes-Hoitinga, and K. Brandt, *Inhibitory effects of feeding with carrots or (−)-falcarinol on development of azoxymethane-induced preneoplastic lesions in the rat colon.* J Agric Food Chem, 2005. 53:1823-7.

40. Bu'Lock, J.D. and G.N. Smith, *The origin of naturally-occurring acetylenes.* J Chem Soc, 1967. 332-6.

41. Dembitsky, V.M., *Anticancer activity of natural and synthetic acetylenic lipids.* Lipids, 2006. 41:883-924.

42. Magnus, V., G. Lacan, R.T. Aplin, and V. Thaller, *Glycerol tridehydrocrepenynate from the basidiomycete Craterellus cornucopioides.* 1989. 28:3047-50.

43. Pang, Z. and O. Sterner, *Cibaric acid, a new fatty acid derivative formed enzymically in damaged fruit bodies of Cantharellus cibarius (Chanterelle).* J Org Chem, 1991. 56:1233-5.

44. Andrade, P.B., B. Ribeiro, P. Valentao, P. Baptista, and R.M. Seabra, *Phenolic compounds, organic acids profiles and antioxidative properties of beefsteak fungus (Fistulina hepatica).* Food Chem Toxicol, 2007. 45:1805-13.

45. Berger, R.G., S.M. Wu, H. Zorn, and U. Krings, *Volatiles from submerged and surface-cultured beefsteak fungus, Fistulina hepatica.* Flavour Fragr J, 2007. 22:53-60.

46. Jones, E.R.H., G. Lowe, and P.V.R. Shannon, *Natural acetylenes. Part XX. Tetra-acetylenic and other metabolites from Fistulina hepatica(Huds) Fr.* J Chem Soc, 1966. 139-44.

47. Tsuge, N., T. Mori, T. Hamano, H. Tanaka, K. Shin-ya, and H. Seto, *Cinnatriacetins A and B, new antibacterial triacetylene derivatives from the fruiting bodies of Fistulina hepatica.* J Antibiot (Tokyo), 1999. 52:578-81.

48. Farrell, I.W., J.W. Keeping, M.G. Pellatt, and V. Thaller, *Natural acetylenes. XLI. Polyacetylenes from fungal fruiting bodies.* J Chem Soc, Perkin 1, 1973. 22:2642-3.

49. Barley, G.C., U. Graf, C.A. Higham, M.Y. Jarrah, E.R.H. Jones, I. O'Neill, R. Tachikawa, V. Thaller, J.L. Turner, and A.V. Hodge, *Natural acetylenes. LXI: Fungal polyacetylenes and the crepenynate pathway. The biosynthesis of some C[9]-C[14] polyacetylenes in fungal cultures.* J Chem Res Synop: Science Reviews, 1987. 1801-38

50. Farrell, I.W., C.A. Higham, E.R.H. Jones, and V. Thaller, *Natural acetylenes. LXII: Fungal polyacetylenes and the crepenynate pathway. Experiments relevant to the biogenesis of the acetylenic bond*. J Chem Res Synop, 1987. 234-5

51. Moller, E.M., G. Bahnweg, H. Sandermann, and H.H. Geiger, *A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues.* Nucleic Acids Res, 1992. 20:6115-6.

52. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method.* Methods Enzymol, 2002. 350:87-96.

53. Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, *Current protocols in molecular biology*. Vol. 2. 1997, New York, NY: Wiley Interscience.

54. Minto, R.E., B.J. Blacklock, H. Younus, and A.C. Pratt, *Atypical biosynthetic properties of a Δ12/v+3 desaturase from the model basidiomycete Phanerochaete chrysosporium.* Appl Environ Microbiol, 2009. 75:1156-64.

55. Christie, W.W., *Mass spectrometry of fatty acids with methylene-interrupted ene-yne systems.* Chem Phys Lipids, 1998. 94:35-41.

56. Marquardt, T.C. and R.F. Wilson, *An improved reversed-phase thin-layer chromatography method for separation of fatty acid methyl esters.* J Am Oil Chem Soc, 1998. 75:1889-92.

57. Buist, P.H., *Fatty acid desaturases: selecting the dehydrogenation channel.* Nat Prod Rep, 2004. 21:249-62.

58. Mikolajczak, K.L., M.O. Bagby, R.B. Bates, and I.A. Wolff, *Prototropic rearrangement of a 1,4-enyne products and mechanism.* J Org Chem, 1965. 30:2983-8.

59. Lamberto, M. and R.G. Ackman, *Positional isomerization of trans-3-hexadecenoic acid employing 2-amino-2-methyl-propanol as a derivatizing agent for ethylenic bond location by gas-chromatography/mass spectrometry.* Anal Biochem, 1995. 230:224-8.

60. Kuklev, D.V. and W.L. Smith, *A procedure for preparing oxazolines of highly unsaturated fatty acids to determine double bond positions by mass spectrometry.* J Lipid Res, 2003. 44:1060-6.

61. Hansen, L. and P.M. Boll, *Polyacetylenes in Araliaceae: their chemistry, biosynthesis and biological significance.* Phytochemistry, 1986. 25:285-93.

62. Nam, J.W. and T.J. Kappock, *Cloning and transcriptional analysis of Crepis alpina fatty acid desaturases affecting the biosynthesis of crepenynic acid.* J Exp Bot, 2007. 58:1421-32.

63. Barnes, J., L.A. Anderson, S. Gibbons, and J.D. Phillipson, *Echinacea species (Echinacea angustifolia (DC.) Hell., Echinacea pallida (Nutt.) Nutt., Echinacea purpurea (L.) Moench): a review of their chemistry, pharmacology and clinical properties.* J Pharm Pharmacol, 2005. 57:929-54.

64. Awang, D.V.C. and D.G. Kindack, *Herbal medicine: Echinacea.* Can Pharm J, 1991. 124:512.

65. Sharma, S.M., M. Anderson, S.R. Schoop, and J.B. Hudson, *Bactericidal and anti-inflammatory properties of a standardized Echinacea extract (Echinaforce): dual actions against respiratory bacteria.* Phytomedicine, 2010. 17:563-8.

66. Senchina, D.S., A.E. Martin, J.E. Buss, and M.L. Kohut, *Effects of Echinacea extracts on macrophage antiviral activities.* Phytother Res, 2010. 24:810-6.

67. Hudson, J. and S. Vimalanathan, *Echinacea—a source of potent antivirals for respiratory virus infections.* Pharmaceuticals, 2011. 4:1019-31.

68. Haller, J., J. Hohmann, and T.F. Freund, *The effect of Echinacea preparations in three laboratory tests of anxiety: comparison with chlordiazepoxide.* Phytother Res, 2010. 24:1605-13.

69. Hajos, N., N. Holderith, B. Nemeth, O.I. Papp, G.G. Szabo, R. Zemankovics, T.F. Freund, and J. Haller, *The effects of an Echinacea preparation on synaptic transmission and the firing properties of CA1 pyramidal cells in the hippocampus.* Phytother Res, 2012. 26:354-62

70. Hohmann, J., D. Redei, P. Forgo, P. Szabo, T.F. Freund, J. Haller, E. Bojnik, and S. Benyhe, *Alkamides and a neolignan from Echinacea purpurea roots and the interaction of alkamides with G-protein-coupled cannabinoid receptors.* Phytochemistry, 2011. 72:1848-53.

71. Woelkart, K. and R. Bauer, *The role of alkamides as an active principle of Echinacea.* Planta Med, 2007. 73:615-23.

72. Christensen, L.P. and J. Lam, *Acetylenes and related compounds in Heliantheae.* Phytochemistry, 1991. 30:11-49.

73. Bohlmann, F. and M. Grenz, *Polyacetylenverbindungen, CXII. Über die Inhaltsstoffe aus Echinacea-Arten.* Chem Ber, 1966. 99:3197-3200.

74. Schulte, K.E., G. Rucker, and J. Perlick, *The presence of polyacetylene compounds in Echinacea purpurea and Echinacea angustifolia DC.* Arzneimittelforschung, 1967. 17:825-9.

75. Beopoulos, A., Z. Mrozova, F. Thevenieau, M.T. Le Dall, I. Hapala, S. Papanikolaou, T. Chardot, and J.M. Nicaud, *Control of lipid accumulation in the yeast Yarrowia lipolytica.* Appl Environ Microbiol, 2008. 74:7779-89.

76. Mauersberger, S. and J. Nicaud, *Tagging of genes by insertional mutagenesis in the yeast Yarrowia lipolytica.* , in *Non-conventional yeasts in genetics, biochemistry and biotechnology*, K. Wolf, K. Breunig, and G. Barth, Editors. 2003, Springer-Verlag: Berlin. p. 343-56.

77. McCartney, A.W., J.M. Dyer, P.K. Dhanoa, P.K. Kim, D.W. Andrews, J.A. McNew, and R.T. Mullen, *Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini.* Plant J, 2004. 37:156-73.

78. Cosson, P., Y. Lefkir, C. Demolliere, and F. Letourneur, *New COP1-binding motifs involved in ER retrieval.* EMBO J, 1998. 17:6863-70.

APPENDICES

Appendix A *Fistulina* Primers, Sequences, and Mass Spectra

Table 14: List of all primers used in *Fistulina hepatica* gene isolation.

Primer Name	Amplification Use	Sequence (From 5' to 3')
Fhe04128-5 RACE 1	5' RACE	CAG TGG TTG GCA AGG ATG TAC GGG AC
Fhe04128-3 RACE 1	5' RACE	CTT TAT TTC GTC CCG TAC ATC CTT GCC
Fh04128s	Full length	ATG CTC TCC ATC TTC GAA GAC AGC C
Fh04128a_NotI	Full length	TTT GCG GCC GCT TAC TCT GGT GTC GGA CGA TTG TTG TC
BamHI-H6(Sc)-Fh04128s	Hexahistidine tagging of full length gene	AAA GGA TCC AAA ATG CAC CAC CAC CAC CAC CAC CTC TCC ATC TTC G
Fh10752-5'C	5' RACE	GTC AGA GTA CGC GTA GTG CTC GCC GAT G
Fh10752-3'C	5' RACE	TCC TCT GGA CAC CGT ACT TCT CAT GGA G
Fh10752ORFs	Full length	ATG GTT TCG ACT AAA GAA TAC CGC
Fh10752-NotI	Full length	GTG GCG GCC GCT TAC TCG GCT TTG GTC TTC TTG AC
BamHi-H6(Sc)-Fh10752s1	Hexahistidine tagging of full length gene	AAA GGA TCC AAA ATG CAT CAT CAT CAT CAT CAT GTT TCG ACT AAA GAA TAC

Nucleotide sequence for Fhe04128-9:

ATGCTCTCCATCTTCGAAGACAGCCCAGAATACAAGAAGCGCGCGCAGACGCCCTTCACGCCGA
CCAAGGTCACTCTCGCTGAAGTACGCGCGGTCGTGCCGAAGCATCTACACGAGAAGAACGCGT
TGAAGGCAACTTTCTATGTCTTACGCGATATTCTCTGCGCGGTCGCGGTCTACAAACTTGGCTGG
CTTATCGATCCTTTGACGACGTCTCTCGTTCGCGACTATGGTTTCAACCCCACTCTTGGCTCAATC
GTGAAATGGGGTCTATGGGCGACGTACTGGCACTGGCAGGGCGTTATTCTTGCCGGCTGGTGG
TGTTTGGGCCACGAAGCTGGACACGGCTCGCTCTCGTCTGCCAAGTGGTACAACACCGTCGTCG
GATATGGTCTGCATACGTTCATCCTCGTTCCTTACTTCTCGTGGCGTTCGTCCCACCATGCGCATC
ATAAGGCCACTATGTCGATCGAACGCGATGAGAACTTCGTTCCCCGCACTCGCAAGGATTATGG
TCTACCCCCGGAAAAGATTGCGGTTCTGAACGATTATCATGAGATTTTTGAAGAGGCGCCATTTT
ACACGCTTGCTCGCATGCTTTTCATGCAGGGTTTGGGTTGGCAATATTACCTTTTCACGAACGCA
CTGGGAAATCCGATGTACCCGCCTGGCACTAACCACTTCAGCCCGTCGTCACCTCTCTTCAAGCC
GCATGAAAGGAATGGCATTATCGCCTCGAACATCGGCCTCACCATCATGAGCAGCCTTCTCTGG
CGTTGGACGATGAGCGTCGGTATCAACATGTTCCTGAAGCTTTATTTCGTCCCGTACATCCTTGC
CAACCACTGGATCGTAATGCTGACGTACCTTCATCACTCCGACCCCACCGTCGCCTACTATCGCG
CGAGCCACTGGTCATTCTTGCGTGGTGCAGTGTCAACTGTTGATCGCCCCCTCCTCGGATGGGC
CGGTCGTTTCTTCCTCCATAACGTCTCGCACGACCACGTTGCCCACCATCTCTTCTCTTCCATCCCT
TTCTACAACCAGCCCGAAGTTACTGAAGCGATCAAGAAGGTGCTGAAGAACGATTATAACTACG
ACTCGACGAACACATTCCGTGCGCTGTATCGCACTTTTTCGCAATGCTGCTTCGTCGAAGATGAT
GGTGATGTTGTCTTCTACAAGAACAGGGACGGCAAGGCAATTCGTATCCTTGCTGAGGGTGCG
CTTGACAACAATCGTCCGACACCAGAGTAA

Translation of coding sequence for Fhe04128-9:

MLSIFEDSPEYKKRAQTPFTPTKVTLAEVRAVVPKHLHEKNALKATFYVLRDILCAVAVYKLGWLIDPL
TTSLVRDYGFNPTLGSIVKWGLWATYWHWQGVILAGWWCLGHEAGHGSLSSAKWYNTVVGYGL
HTFILVPYFSWRSSHHAHHKATMSIERDENFVPRTRKDYGLPPEKIAVLNDYHEIFEEAPFYTLARML
FMQGLGWQYYLFTNALGNPMYPPGTNHFSPSSPLFKPHERNGIIASNIGLTIMSSLLWRWTMSVGI
NMFLKLYFVPYILANHWIVMLTYLHHSDPTVAYYRASHWSFLRGAVSTVDRPLLGWAGRFFLHNVS
HDHVAHHLFSSIPFYNQPEVTEAIKKVLKNDYNYDSTNTFRALYRTFSQCCFVEDDGDVVFYKNRDG
KAIRILAEGALDNNRPTPE.

Nucleotide sequence for Fhe04128-10:

ATGCTCTCCATCTTCGAAGACAGCCCAGAATACAAGAAGCGCGCGCAGACGCCCTTCACGCCGA
CCAAGGTCACTCTCGCTGAAGTACGCGCGGTCGTGCCGAAGCATCTACACGAGAAGAACGCGT
TGAAGGCAACTTTCTATGTCTTACGCGATATTCTCTGTGCGGTCGCGGTCTACAAACTTGGCTGG
CTTATCGATCCTTTGACGACGTCTCTCGTTCGCGACTATGGTTTCAACCAGACTCTTGGCTCAATC
GTGAAATGGGGTCTATGGGCGACGTACTGGCACTGGCAGGGCGTTATTCTTGCCGGCTGGTGG
TGCTTGGGCCACGAAGCTGGACACGGCTCGCTCTCGTCTGCCAAGTGGTACAACACCGTCGTCG
GATATGGTCTGCATACGTTCATCCTCGTTCCTTACTTCTCGTGGCGTTCGTCCCACCATGCGCATC
ATAAGGCCACTATGTCGATCGAACGCGATGAGAACTTCGTTCCCCGCACTCGCAAGGATTATGG
TCTACCCCCGGAAAAGATTGCGGTTCTGAACGATTATCATGAGATTTTTGAAGAGGCGCCATTTT
ACACGCTTGCCCGCATGCTTTTCATGCAGGGTTTGGGTTGGCAATATTACCTTTTCACGAACGCA
CTGGGAAATCCGATGTACCCGCCTGGCACTAACCACTTCAGCCCGTCGTCACCTCTCTTCAAGCC
GCATGAAAGGAATGGCATTATCGCCTCGAACATCGGCCTCACCATCATGAGCAGCATTCTCTGG
CGTTGGACGATGAGCGTCGGTGTCAACATGTTCCTGAAGCTTTATTTCGTCCCGTACATCCTTGC
CAACCACTGGATCGTAATGCTAACGTACCTTCATCACTCCGACCCCACCGTCGCCTATTATCGCG
CGAGCCACTGGTCGTTCTTGCGTGGTGCAGTGTCAACTGTTGATCGCCCCCTTCTCGGATGGGC
CGGTCGTTTCTTCCTCCATAACGTCTCGCACGACCACGTTGCGCACCATCTCTTCTCTTCTATCCC
TTTCTACAACCAGCCCGAAGTTACTGAAGCGATCAAGAAGGTGCTGAAGAACGATTATAACTAC
GACTCGACGAACACATTCCGTGCGCTGTATCGCACTTTTTCGCAATGCTGCTTCGTTGAAGATGA
CGGTGATGTTGTCTTCTACAAGAACAGGGACGGCAAGGCAATTCGTGTCCTTGCTGAGGGTGC
GCTTGACAACAATCGTCCGACACCAGAGTAA

Translation of coding sequence for Fhe04128-10:

MLSIFEDSPEYKKRAQTPFTPTKVTLAEVRAVVPKHLHEKNALKATFYVLRDILCAVAVYKLGWLIDPL
TTSLVRDYGFNQTLGSIVKWGLWATYWHWQGVILAGWWCLGHEAGHGSLSSAKWYNTVVGYG
LHTFILVPYFSWRSSHHAHHKATMSIERDENFVPRTRKDYGLPPEKIAVLNDYHEIFEEAPFYTLARM
LFMQGLGWQYYLFTNALGNPMYPPGTNHFSPSSPLFKPHERNGIIASNIGLTIMSSILWRWTMSVG
VNMFLKLYFVPYILANHWIVMLTYLHHSDPTVAYYRASHWSFLRGAVSTVDRPLLGWAGRFFLHNV
SHDHVAHHLFSSIPFYNQPEVTEAIKKVLKNDYNYDSTNTFRALYRTFSQCCFVEDDGDVVFYKNRD
GKAIRVLAEGALDNNRPTPE.

Nucleotide sequence for Fhe10752-2:

ATGGTTTCGACTAAAGAATACCGCGAAGAGGACCTACCCGCTTATACCCCAATGCCCTGGTCTC
TGAAGGAGATCAGGGATGCCATCCCGCCCCAATACTTTGTCAAGGACACCCCTCGGGGCTTGA
GTTATCTGGCGCGGGATGTTGTCATGGCCGCCGCTTTGTGGTACGGTGCTACGTTCATTGACCC
GTACTTCAAGAGCGCAGAGGCTGCGAGCGTCCTCACGCCCGTTGGTGCTCAGGTGGCCCGTTG
GGCTAGCTGGGGTGTCTACTGGTGGTTCCAGGGTCTTGTCTTCACCGGCATTTGGGTGATTGGT
CACGAGTGCGGACACGGTGCGTTCTCGGACCACAAGATCGTCAACGATGTCGTTGGTTTTGTCA
CGCACACTCTCCTCTGGACACCGTACTTCTCATGGAGAATCTCGCACCACCGTCATCACTCGAAT
CACGCCTCCATGGAGCGCGACGAGGTGTACGTTCCGAAGACGCGCGAGGACCTCGGCATTCCG
GATGAGAAGGAAGGTGTCAAGATCGACTACGAGGAATATTTCGGCGACACGCCCATTTACACG
ATGTACATGCTCATTCGTCAGCAGATCCTCGCCTTCCCTGCCTATCTGCTCTTCAACGTCTCTGGA
CAAAAACACTATCCCAAGTGGACTAACCACTTTGATCCAAACTCGATCCTTTTCACCAAGTCGCA
GCGCAATGTTGTCATTATCAGCAACATCGGTATCGCGGCCATGGTTTACGGCGTCATGTACTTCA
GCAACGCCTACTCGTTCGCCCAGGTCTGCAAGTACTACTTCCTCCCGTGGCTAGAGGTGACTCAC
TGGTTCATCATGATCACCTATCTCCACCACACGTCGCCCGACCTCCCCCACTACCGTGGCAAGGA
GTGGAGCTTCCAGCGTGGCGCTGCTGCCACTGTTGACCGCAACTTCCTCGGCTGGCAAGGCCGC
TTCTTCCTCCATGATGTCGCGCACTACCACGTCATCCACCACTTCTTCCCCAAGATGCCGTTCTAT
CACGGTCCTGAGGCAACAAAGTATCTCAAGGACTTCATCGGCGAGCACTACGCGTACTCTGACC
AGCCTGTCTTCACCGCCCTCTGGGAAAGCTATAATAAGTGTCAGTTCGTCGAGAACGAGGGTGA
TGTCCTTTTCTACCGCGACTCCGAAGGCAAGGCTGTTCGACGTCCCGCTGACCAGTACATGGTC
AAGAAGACCAAAGCCGAGTAA

Translation of coding sequence for Fhe10752-2:

MVSTKEYREEDLPAYTPMPWSLKEIRDAIPPQYFVKDTPRGLSYLARDVVMAAALWYGATFIDPYFK
SAEAASVLTPVGAQVARWASWGVYWWFQGLVFTGIWVIGHECGHGAFSDHKIVNDVVGFVTHT
LLWTPYFSWRISHHRHHSNHASMERDEVYVPKTREDLGIPDEKEGVKIDYEEYFGDTPIYTMYMLIR
QQILAFPAYLLFNVSGQKHYPKWTNHFDPNSILFTKSQRNVVIISNIGIAAMVYGVMYFSNAYSFAQ
VCKYYFLPWLEVTHWFIMITYLHHTSPDLPHYRGKEWSFQRGAAATVDRNFLGWQGRFFLHDVAH
YHVIHHFFPKMPFYHGPEATKYLKDFIGEHYAYSDQPVFTALWESYNKCQFVENEGDVLFYRDSEGK
AVRRPADQYMVKKTKAE.

Nucleotide sequence for Fhe10752-7:

ATGGTTTCGACTAAAGAATACCGCGAAGAGGACCTACCCGCTTATACCCCAATGCCCTGGTCTC
TGAAGGAGATCAGGGATGCCATCCCGCCCCAATACTTTGTCAAGGACACCCCTCGGGGCTTGA
GTTATCTGGCGCGGGATGTTGTCATGGCCGCCGCTTTGTGGTACGGTGCTACGTTCATTGACCC
GTACTTCAAGAGCGCAGAGGCTGCGAGCGTCCTCACGCCCGTTGGTGCTCAGGTGGCCCGTTG
GGCTAGCTGGGGTGTCTACTGGTGGTTCCAGGGTCTTGTCTTCACCGGCATTTGGGTGATTGGT
CACGAGTGCGGACACGGTGCGTTCTCGGACCACAAGATCGTCAACGATGTCGTTGGTTTTGTCA
CGCACACTCTCCTCTGGACACCGTACTTCTCATGGAGAATCTCGCACCACCGTCATCACTCGAAT
CACGCCTCCATGGAGCGCGACGAGGTGTACGTTCCGAAGACGCGCGAGGACCTCGGCATTCCG
GATGAGAAGGAAGGTGTCAAGATCGACTACGAGGAATATTTCGGCGACACGCCCATTTACACG
ATGTACATGCTCATTCGTCAGCAGATCCTCGCCTTCCCTGCCTATCTGCTCTTCAACGTCTCTGGA
CAAAAACACTATCCCAAGTGGACTAACCACTTTGATCCAAACTCGATCCTTTTCACCAAGTCGCA
GCGCAATGTTGTCATTATCAGCAACATCGGTATCGCGGCCATGGTTTACGGCGTCATGTACTTCA
GCAACGCCTACTCGTTCGCCCAGGTCTGCAAGTACTACTTCCTCCCGTGGCTAGAGGTGACTCAC
TGGTTCATCATGATCACCTATCTCCACCACACGGCACCCGACCTCCCCCACTACCGTGGCAAGGA
GTGGAGCTTCCAGCGTGGCGCTGCTGCCACTGTTGACCGCAACTTCCTCGGCTGGCAAGGCCGC
TTCTTCCTCCATGATGTCGCGCACTACCACGTCATCCACCACTTCTTCCCCAAGATGCCGTTCTAT
CACGGTCCTGAGGCAACAAAGTATCTCAAGGACTTCATCGGCGAGCACTACGCGTACTCTGACC
AGCCTGTCTTCGCCGCCCTCTGGGAAAGCTATAATAAGTGTCAGTTCGTCGAGAACGAGGGTGA
TGTCCTTTTCTACCGCGACTCCGAAGGCAAGGCTGTTCGACGTCCCGCTGACCAGTACATGGTC
AAGAAGACCAAAGCCGAGTAA

Translation of coding sequence for Fhe10752-7:

MVSTKEYREEDLPAYTPMPWSLKEIRDAIPPQYFVKDTPRGLSYLARDVVMAAALWYGATFIDPYFK
SAEAASVLTPVGAQVARWASWGVYWWFQGLVFTGIWVIGHECGHGAFSDHKIVNDVVGFVTHT
LLWTPYFSWRISHHRHHSNHASMERDEVYVPKTREDLGIPDEKEGVKIDYEEYFGDTPIYTMYMLIR
QQILAFPAYLLFNVSGQKHYPKWTNHFDPNSILFTKSQRNVVIISNIGIAAMVYGVMYFSNAYSFAQ
VCKYYFLPWLEVTHWFIMITYLHHTAPDLPHYRGKEWSFQRGAAATVDRNFLGWQGRFFLHDVA
HYHVIHHFFPKMPFYHGPEATKYLKDFIGEHYAYSDQPVFAALWESYNKCQFVENEGDVLFYRDSE
GKAVRRPADQYMVKKTKAE.

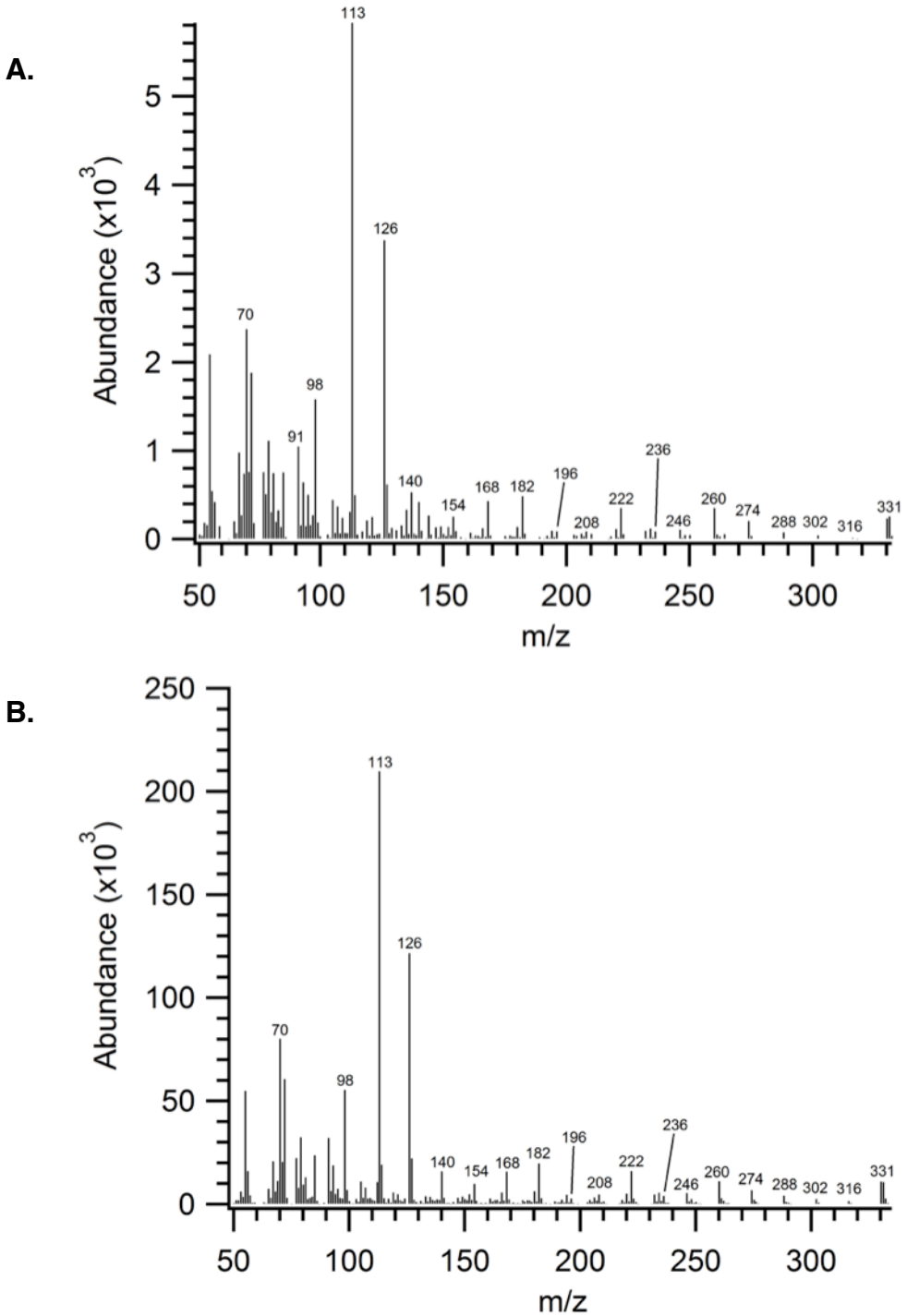


Figure 41: Mass spectral analysis of crepenynic acid pyrrolidide. Shown are mass spectra of (A) pyrrolidides synthesized from an expression culture of Fhe10752 and (B) the authentic standard purified from *C. alpina* seed oil.

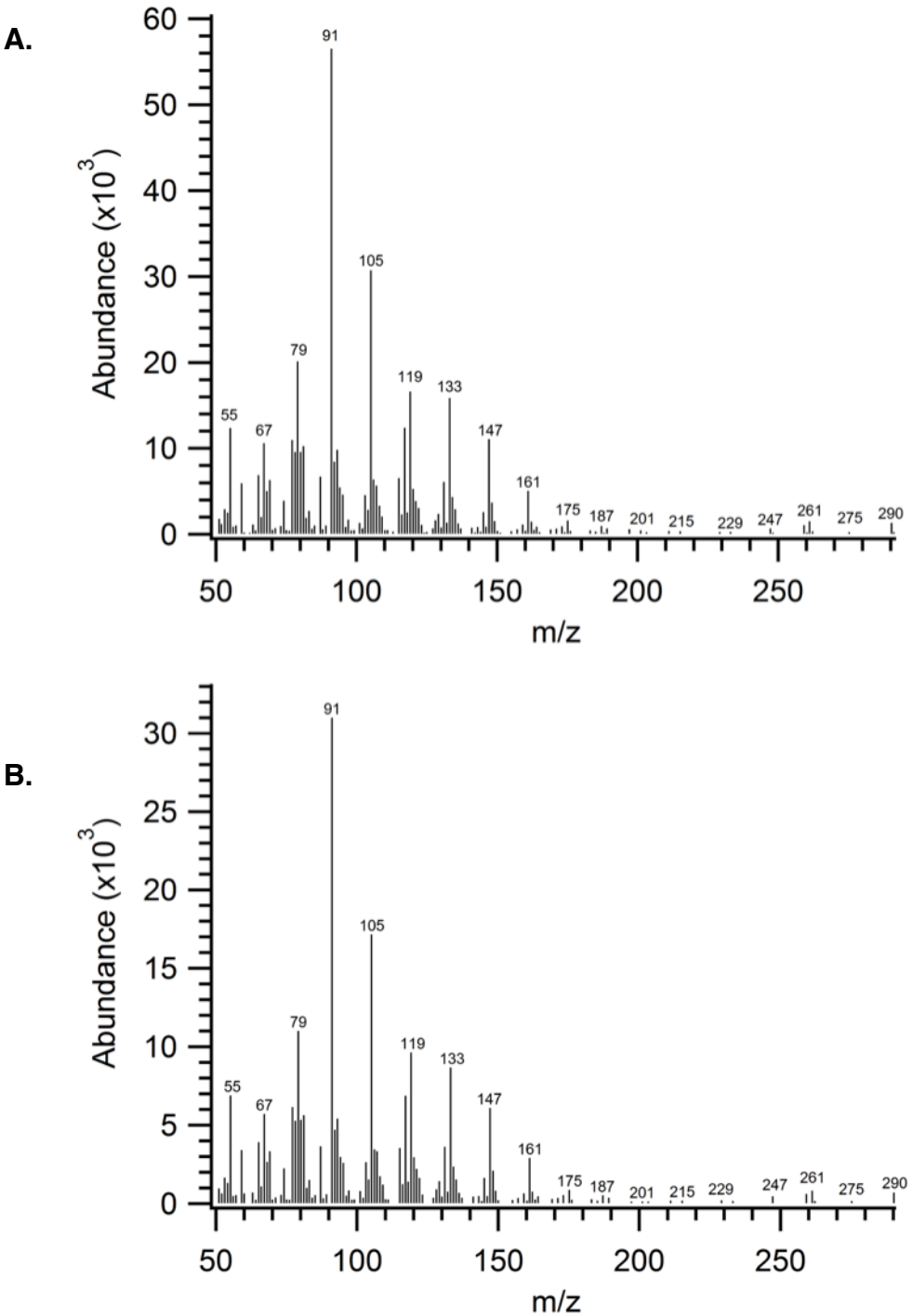


Figure 42: Stacked comparison of mass spectra of methyl dehydrocrepenynate. Provided is the (A) mass spectrum obtained from FAMEs prepared from Fhe04128 expression cultures supplemented with crepenynate and (B) from an authentic standard resulting form the derivatization of *Cantharellus* fruit bodies.

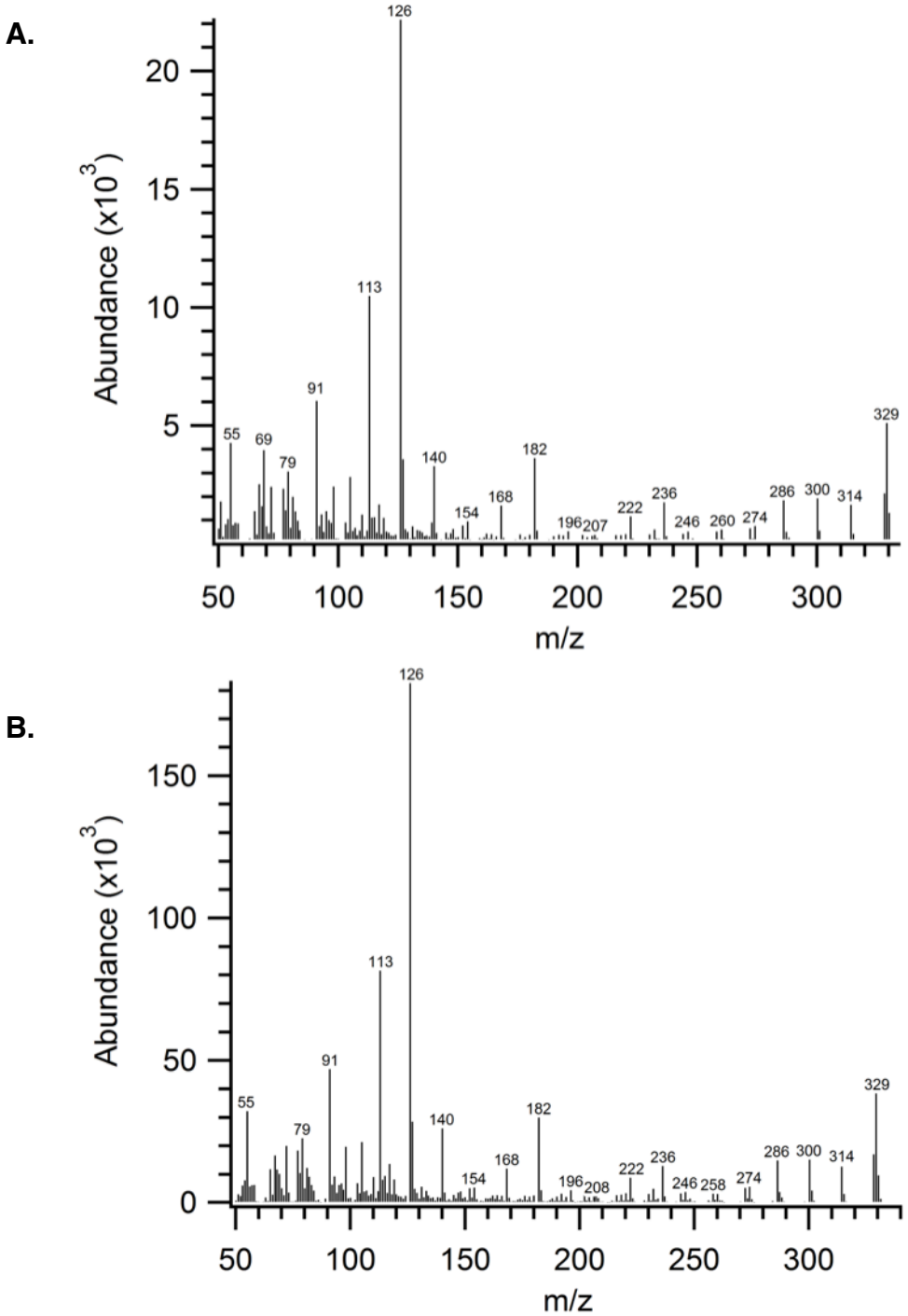


Figure 43: Comparison of GC-MS data for dehydrocrepenynoyl DMOX derivatives. Mass spectra corresponding to (A) dehydrocrepenynoyl DMOX from 18:1$^{\Delta 9c,12a}$-supplemented Fhe04128-expressing yeast and (B) dehydrocrepenynoyl DMOX prepared from authentic *C. cibarius* standard.

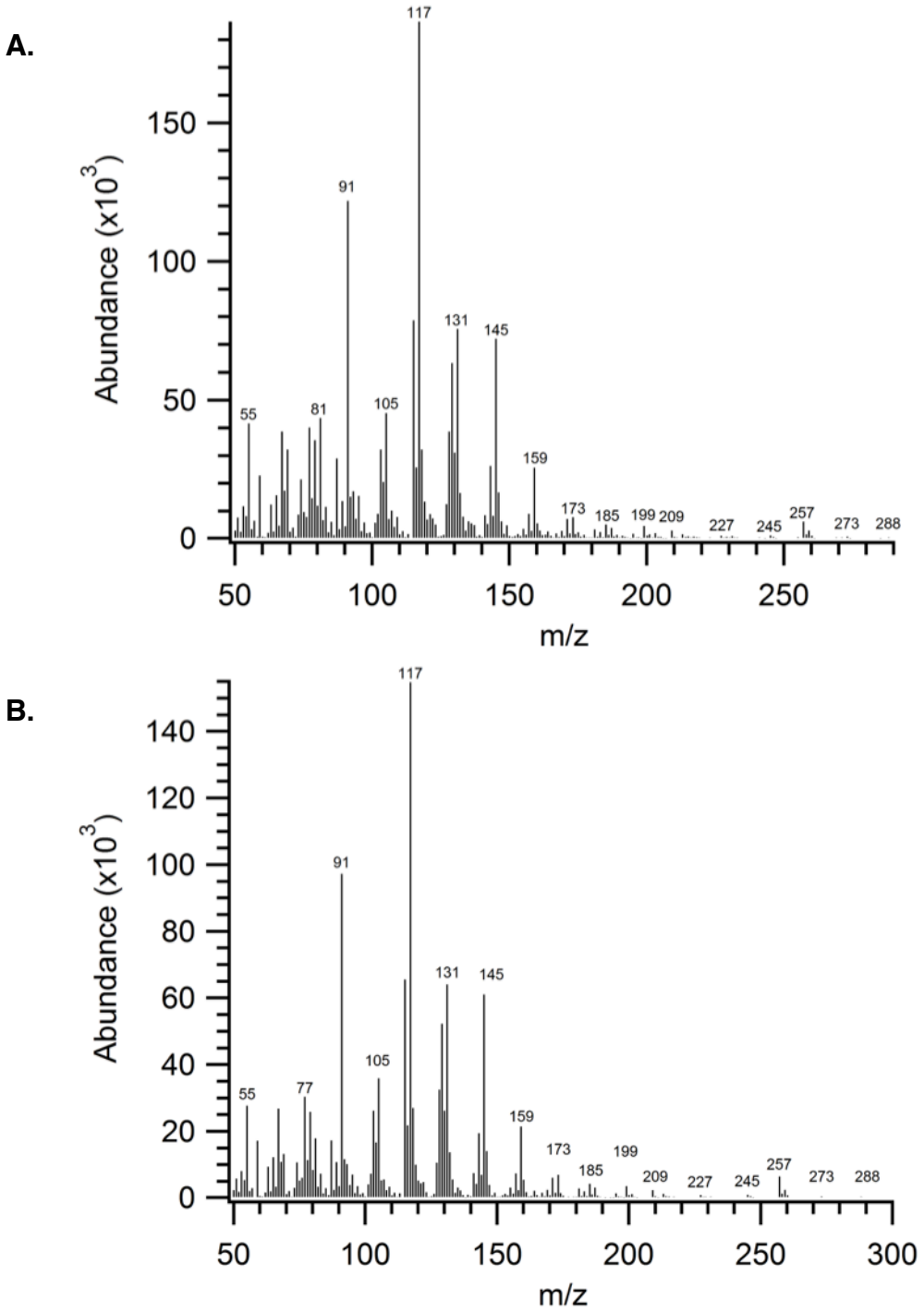


Figure 44: GC-MS analysis of 18:1$^{\Delta 9c,12a,14a}$ FAMEs. (A) Mass spectrum of the diacetylenic FAME developed from Fhe04128 expression cultures supplemented with crepenynic acid. (B) Mass spectrum of synthetic standard FAME prepared in our lab.

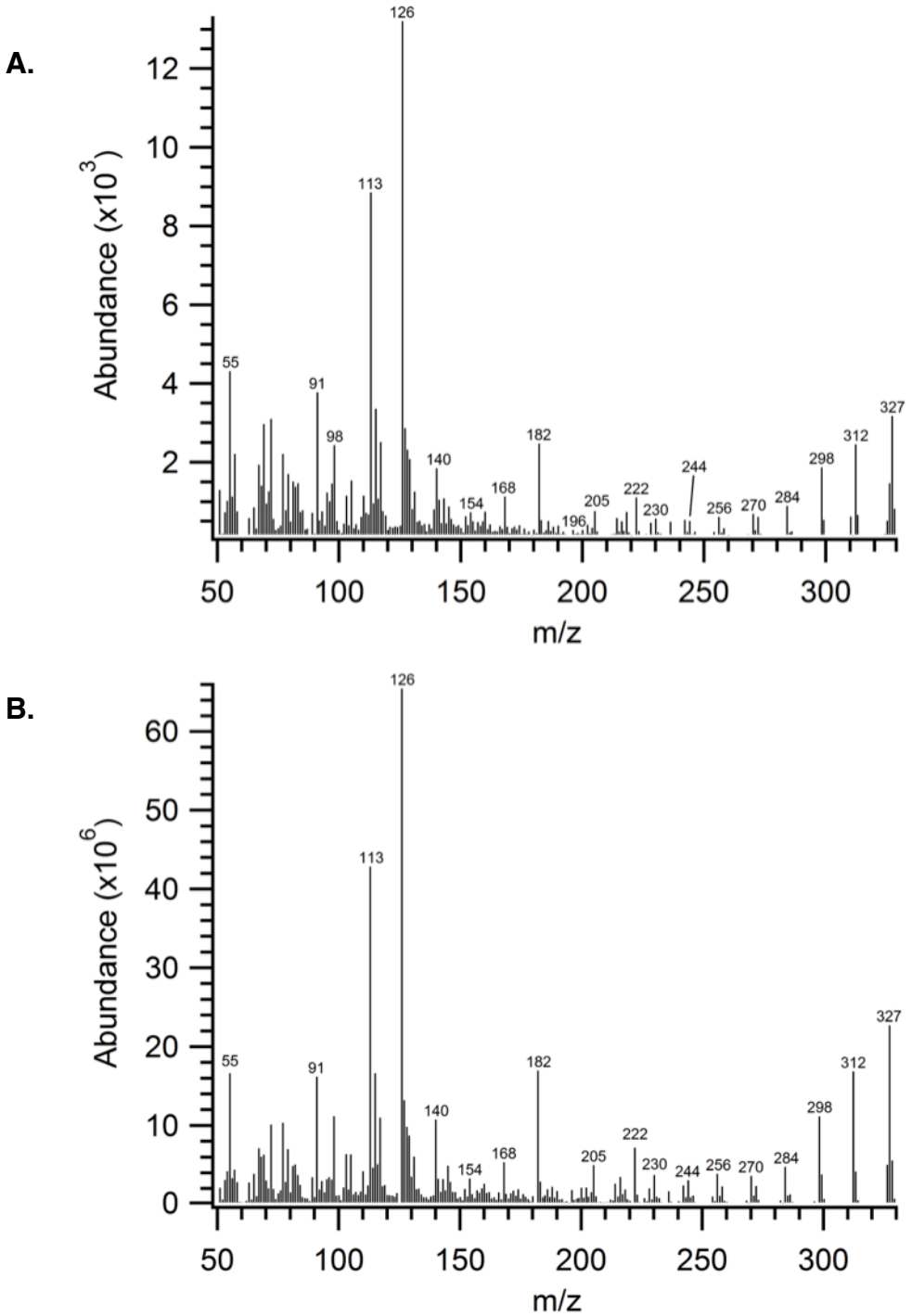


Figure 45: Mass spectral identification of 18:1$^{\Delta 9c,12a,14a}$ dimethyloxazolines. (A) Mass spectrum for the diacetylenic DMOX derivative prepared from transgenic yeast expressing Fhe04128 supplemented with crepenynic acid. (B) Mass spectrum for diacetylenic DMOX derivative produced from the synthetic FAME standard.

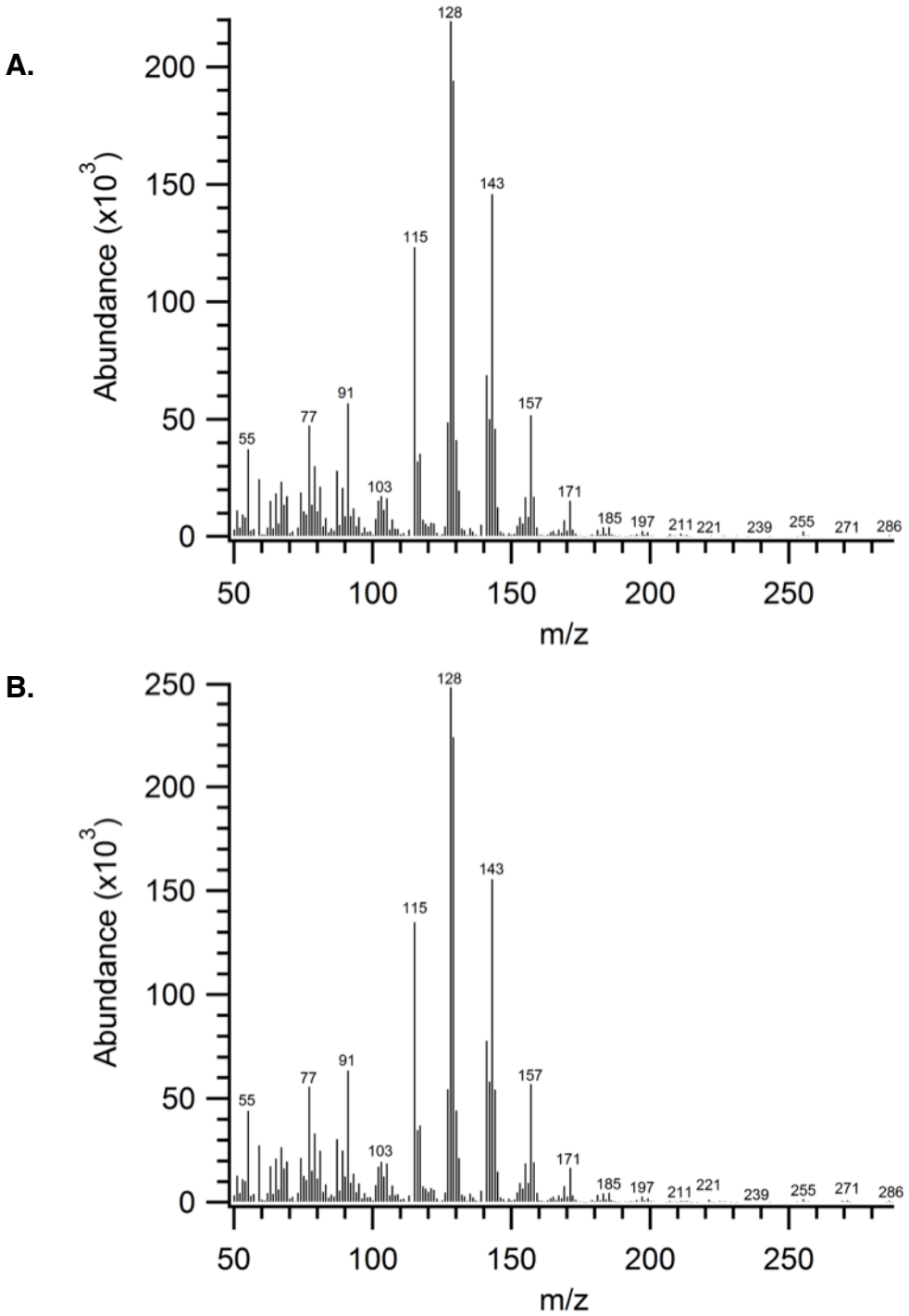


Figure 46: Mass fragmentation data for methyl 18:2$^{\Delta 9c,12a,14a,16}$ isomers. (A) (16*Z*)-dehydrodiacetylenic and (B) (16*E*)-dehydrodiacetylenic fatty acid methyl esters prepared from Fhe04128 transgenic yeast supplemented with dehydrocrepenynate.

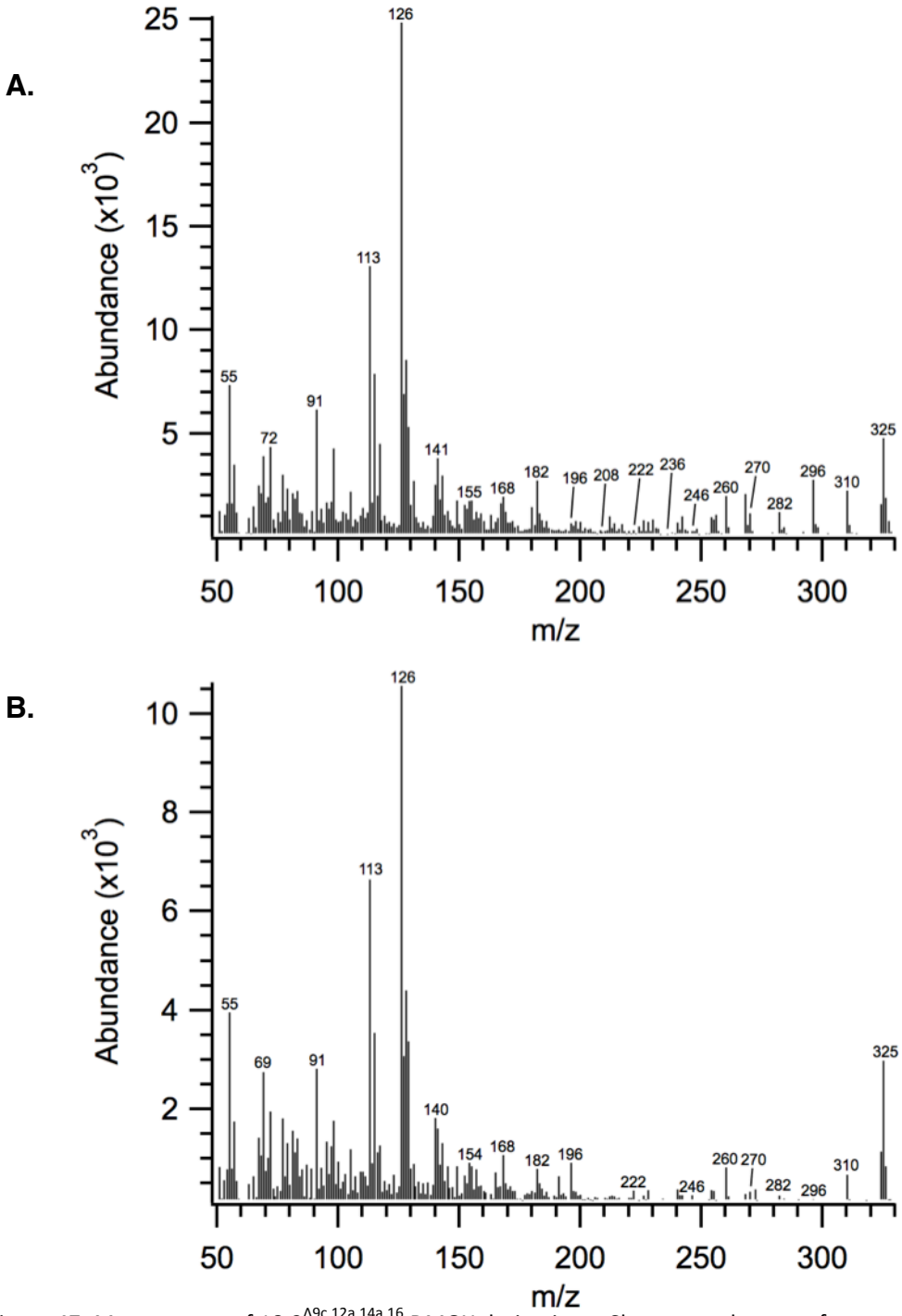


Figure 47: Mass spectra of 18:2$^{\Delta 9c,12a,14a,16}$ DMOX derivatives. Shown are the mass fragmentation patterns of (A) 18:2$^{\Delta 9c,12a,14a,16c}$ and (B) 18:2$^{\Delta 9c,12a,14a,16t}$ made from Fhe04128 expressing-yeast supplemented with dehydrocrepenynate.

Appendix B *Echinacea* Primers, Sequences, and Mass Spectra

Table 15: A comprehensive list of primers used for the isolation of *Echinacea purpurea* gene targets.

Primer Name	Amplification Use	Sequence (From 5' to 3')
Epa955-5'A	5' RACE	TGA GAT CAG ATG ATG CAA GAC ATG AGT A
Epa955-3'A	3' RACE	GAA ATA CAG TCA TCG TAA TCA CCA TGC C
Epa955-5'B	5' RACE	GCT CAC CCA ACA CCG GCT TGA TGG CGT C
Epa955-3'B	3' RACE	GTT TTG TCA AGG TAG TAT CCT AAC AGG T
Epa955s (RACE)	Full length	ATG GGT GCG GGT GGC CGG ATG ACC A
Epa955s (Illum)	Full length	ATG GGT GCT GGT GGG CGG ATG ACC A
Epa955-NotI	Full length	GTG GCG GCC GCT TAC AAC TTG TGA TCC CAG T
BamHI-H6(Sc)-Epa955 (RACE)	Hexahistidine tagging of full length gene	AAA GGA TCC AAA ATG CAT CAT CAT CAT CAT CAT GGT GCG GGT GGC CGG ATG AC
BamHI-H6-Epa955	Hexahistidine tagging of full length gene	AAA GGA TCC AAA ATG CAC CAC CAC CAC CAC CAC GGT GCT GGT GGG CGG ATG AC
YlHis6Epa955repair_s	Site-directed mutagenesis	GCC AAC ACC AAC TCT CTC GAT AAC GAT GAG GTT TAC AT
YlHis6Epa955repair_a	Site-directed mutagenesis	ATG TAA ACC TCA TCG TTA TCG AGA GAG TTG GTG TTG GC
Epa2161-5'A	5' RACE	ACG CCT AGC ACA GGA ACT CCA TAC ATG C
Epa2161-3'A	3' RACE	ACG CTA AGT GAT CTA AAG AAA GCT ATC C
Epa2161s	Full length	ATG GGC GCA GGT GGT CGC ATG TCG G
Epa2161-NotI	Full length	GTG GCG GCC GCT TAC ATT TTA TGG TAC CAA TAG
BamHI-H6-Epa2161	Hexahistidine tagging of full length gene	AAA GGA TCC AAC ATG CAC CAC CAC CAC CAC CAC GGC GCA GGT GGT CGC ATG TC
YlHis6Epa2161repair_s	Site-directed mutagenesis	CCC GGT TGA TCC ACC GTT CAC GCT AAG TGA TCT AAA G
YlHis6Epa2161repair_a	Site-directed mutagenesis	CTT TAG ATC ACT TAG CGT GAA CGG TGG ATC AAC CGG G

Nucleotide sequence for Epa955RACE:

ATGGGTGCGGGTGGCCGGATGACCACTGCCTCCGATGACCAGAAGAACATCTTGGAACGAGTA
CCTTGCTCGAAACCGCCATTTTCATTGGCCGATCTCAAGAAAGCAATTCCGCCTCATTGCTTCGA
AAGATCGGTCATCCGCTCCTCGTACTACGTCGTCCATGACCTTGCTGTTGCCTACGTGTTTTACTA
CCTTGCAAACACATACATCCCACTTCTTCCATCCCCTTTTGCATACTTGGCATGGCCCCTATATTG
GTTTTGTCAAGGTAGTATCCTAACAGGTGTTTGGGTCATTGGTCATGAATGCGGCCACCATGCC
TTCAGTGAGTACCAGTGGATCGATGACACTGTCGGATTTGTCCTACACTCAGCACTCTTCACCCC
GTATTTCTCTTGGAAATACAGTCATCGTAATCACCATGCCAACCCCAACTCTCTCGATAACGATG
AGGGTTACATCCCCAAAGTTAAATCCAAAGTTACACTTTATTCCAAAATCCTTAACAACCCTCCTG
GGCGAGTTTTCACCTTGGCTTTCAGATTGATCCTTGGTTTTCCTTTATACCTTTTCACCAATGTGT
CTGGCAAGAAATACGAGCGTTTTGCAAACCACTTTGATCCAATGAGCCCAATCTTTACTGAACGT
GAACGTGTTCAAGTCTTGCTTTCTGACTTGGGACTCATTGCAGTGGCTTATGCGGTTCGCAAAGC
TGTATTGGCTAAAGGCGCTGCATGGGTTATGTGCATTTACGGAGTTCCTGTTCTTGTTGTAAGTG
CATTCTTTGTTTTGATCACTTATCTACACCACACCCATCTCTCACTGCCTCATTATGACTCAACGGA
GTGGGACTGGATCAAGGGAGCTTTGTCCACCATCGACAGAGATTTCGGGTTCCTGAATAGGGT
TTTCCATGACGTCACGCATACTCATGTCTTGCATCATCTGATCTCATACATCCCACATTATCATGC
TAAGGAGGCTAGGGACGCCATCAAGCCGGTGTTGGGTGAGCATTACAAGATTGACAGGACACC
CATCATCAAGGCAATGTGGAGAGAGGCTAAGGAATGCATCTATATCGAGCCGGATCAAGGTAG
CGAGCACCAAGGTGTGTACTGGTATCACAAGTTGTAA

Translation of coding sequence for Epa955RACE:

MGAGGRMTTASDDQKNILERVPCSKPPFSLADLKKAIPPHCFERSVIRSSYYVVHDLAVAYVFYYLA
NTYIPLLPSPFAYLAWPLYWFCQGSILTGVWVIGHECGHHAFSEYQWIDDTVGFVLHSALFTPYFSW
KYSHRNHHANPNSLDNDEGYIPKVKSKVTLYSKILNNPPGRVFTLAFRLILGFPLYLFTNVSGKKYERF
ANHFDPMSPIFTERERVQVLLSDLGLIAVAYAVRKAVLAKGAAWVMCIYGVPVLVVSAFFVLITYLH
HTHLSLPHYDSTEWDWIKGALSTIDRDFGFLNRVFHDVTHTHVLHHLISYIPHYHAKEARDAIKPVLG
EHYKIDRTPIIKAMWREAKECIYIEPDQGSEHQGVYWYHKL.

Nucleotide sequence for Epa955Trans:

ATGGGTGCGGGTGGCCGGATGACCACTGCCTCCGATGACCAGAAGAACATCTTGGAACGAGTA
CCTTGCTCGAAACCGCCATTTTCATTGGCCGATCTCAAGAAAGCAATTCCGCCTCATTGCTTCGA
AAGATCGGTCATCCGCTCCTCGTACTACGTCGTCCATGACCTTGCTGTTGCCTACGTGTTTTACTA
CCTTGCAAACACATACATCCCACTTCTTCCATCCCCTTTTGCATACTTGGCATGGCCCCTATATTG
GTTTTGTCAAGGTAGTATCCTAACAGGTGTTTGGGTCATTGGTCATGAATGCGGCCACCATGCC
TTCAGTGAGTACCAGTGGATCGATGACACTGTCGGATTTGTCCTACACTCAGCACTCTTCACCCC
GTATTTCTCTTGGAAATACAGTCATCGTAATCACCATGCCAACACCAACTCTCTCGATAACGATG
AGGTTTACATCCCCAAAGTTAAATCCAAAGTTACACTTTATTCCAAAATCCTTAACAACCCTCCTG
GGCGAGTTTTCACCTTGGCTTTCAGATTGATCCTTGGTTTTCCTTTATACCTTTTCACCAATGTGT
CTGGCAAGAAATACGAGCGTTTTGCAAACCACTTTGATCCAATGAGCCCAATCTTTACTGAACGT
GAACGTGTTCAAGTCTTGCTTTCTGACTTGGGACTCATTGCAGTGGCTTATGCGGTTCGCAAAGC
TGTATTGGCTAAAGGCGCTGCATGGGTTATGTGCATTTACGGAGTTCCTGTTCTTGTTGTAAGTG
CATTCTTTGTTTTGATCACTTATCTACACCACACCCATCTCTCACTGCCTCATTATGACTCAACGGA
GTGGGACTGGATCAAGGGAGCTTTGTCCACCATCGACAGAGATTTCGGGTTCCTGAATAGGGT
TTTCCATGACGTCACGCATACTCATGTCTTGCATCATCTGATCTCATACATCCCACATTATCATGC
TAAGGAGGCTAGGGACGCCATCAAGCCGGTGTTGGGTGAGCATTACAAGATTGACAGGACACC
CATCATCAAGGCAATGTGGAGAGAGGCTAAGGAATGCATCTATATCGAGCCGGATCAAGGTAG
CGAGCACCAAGGTGTGTACTGGTATCACAAGTTGTAA

Translation of coding sequence for Epa955Trans:

MGAGGRMTTASDDQKNILERVPCSKPPFSLADLKKAIPPHCFERSVIRSSYYVVHDLAVAYVFYYLA
NTYIPLLPSPFAYLAWPLYWFCQGSILTGVWVIGHECGHHAFSEYQWIDDTVGFVLHSALFTPYFSW
KYSHRNHHANTNSLDNDEVYIPKVKSKVTLYSKILNNPPGRVFTLAFRLILGFPLYLFTNVSGKKYERF
ANHFDPMSPIFTERERVQVLLSDLGLIAVAYAVRKAVLAKGAAWVMCIYGVPVLVVSAFFVLITYLH
HTHLSLPHYDSTEWDWIKGALSTIDRDFGFLNRVFHDVTHTHVLHHLISYIPHYHAKEARDAIKPVLG
EHYKIDRTPIIKAMWREAKECIYIEPDQGSEHQGVYWYHKL.

Nucleotide sequence for Epa955RACE-opt for *Saccharomyces cerevisiae:*

ATGGGTGCTGGTGGTAGAATGACTACAGCATCCGATGATCAAAAGAATATTTTGGAAAGAGTT
CCATGTTCTAAGCCACCTTTTTCATTGGCTGATTTGAAGAAAGCAATTCCACCTCATTGTTTCGAA
AGATCAGTTATTAGATCTTCATACTACGTTGTTCATGATTTGGCTGTTGCATACGTTTTCTATTAC
TTAGCTAACACATACATTCCATTGTTACCATCCCCTTTTGCATATTTGGCTTGGCCTTTATACTGGT
TCTGTCAAGGTAGTATTTTAACCGGTGTTTGGGTTATTGGTCATGAATGTGGTCATCATGCTTTT
TCTGAATACCAATGGATCGATGATACTGTTGGTTTCGTTTTGCATTCAGCTTTGTTTACTCCATAT
TTCTCCTGGAAGTACAGTCATAGAAACCATCATGCAAACACAAACTCTTTGGATAACGATGAAG
TTTACATCCCTAAGGTTAAGTCCAAGGTTACATTGTACAGTAAGATCTTGAACAACCCACCTGGT
AGAGTTTTTACCTTGGCTTTCAGATTGATCTTGGGTTTCCCATTGTATTTGTTTACTAACGTTTCT
GGTAAAAAGTACGAAAGATTCGCAAACCATTTCGATCCAATGTCCCCTATTTTTACTGAAAGAG
AAAGAGTTCAAGTTTTGTTGAGTGATTTGGGTTTGATTGCTGTTGCATATGCTGTTAGAAAGGC
AGTTTTGGCTAAAGGTGCTGCATGGGTTATGTGTATCTATGGTGTTCCAGTTTTGGTTGTTTCAG
CTTTCTTTGTTTTGATCACATATTTGCATCATACCCATTTGTCTTTACCTCATTACGATTCAACCGA
ATGGGATTGGATTAAAGGTGCATTGTCTACTATCGATAGAGATTTCGGTTTCTTAAACAGAGTTT
TCCATGATGTTACCCATACTCATGTTTTGCATCATTTGATCTCATACATTCCACATTACCATGCAA
AGGAAGCTAGAGATGCAATTAAACCAGTTTTGGGTGAACATTACAAGATCGATAGAACACCTAT
TATTAAGGCTATGTGGAGAGAAGCAAAGGAATGTATCTATATCGAACCTGATCAAGGTTCTGAA
CATCAAGGTGTTTACTGGTATCATAAATTATAA

Translation of coding sequence for Epa955RACE-opt for *Saccharomyces cerevisiae:*

MGAGGRMTTASDDQKNILERVPCSKPPFSLADLKKAIPPHCFERSVIRSSYYVVHDLAVAYVFYYLA
NTYIPLLPSPFAYLAWPLYWFCQGSILTGVWVIGHECGHHAFSEYQWIDDTVGFVLHSALFTPYFSW
KYSHRNHHANTNSLDNDEVYIPKVKSKVTLYSKILNNPPGRVFTLAFRLILGFPLYLFTNVSGKKYERF
ANHFDPMSPIFTERERVQVLLSDLGLIAVAYAVRKAVLAKGAAWVMCIYGVPVLVVSAFFVLITYLH
HTHLSLPHYDSTEWDWIKGALSTIDRDFGFLNRVFHDVTHTHVLHHLISYIPHYHAKEARDAIKPVLG
EHYKIDRTPIIKAMWREAKECIYIEPDQGSEHQGVYWYHKL.

Nucleotide sequence for Epa955RACE-opt for *Yarrowia lipolytica:*

ATGGGTGCCGGCGGACGAATGACCACTGCTTCTGACGACCAGAAGAACATCCTGGAGCGAGTC
CCCTGTTCCAAGCCCCCTTTCTCGCTGGCCGACCTCAAGAAGGCTATCCCCCCTCACTGCTTTGA
GCGATCCGTGATTCGATCTTCCTACTACGTGGTCCACGACCTGGCCGTCGCTTACGTTTTCTACT
ACCTCGCCAACACTTACATCCCCCTGCTCCCCTCGCCTTTCGCTTACCTGGCCTGGCCTCTCTACT
GGTTTTGTCAGGGATCTATCCTGACCGGTGTCTGGGTTATTGGTCACGAGTGCGGCCACCATGC
CTTCTCTGAGTACCAGTGGATTGACGACACCGTGGGTTTTGTCCTGCACTCCGCCCTCTTCACTC
CCTACTTTTCGTGGAAGTACTCTCATCGAAACCACCATGCTAACACCAACTCGCTGGACAACGAC
GAGGTCTACATCCCTAAGGTTAAGTCTAAGGTGACCCTGTACTCCAAGATTCTCAACAACCCCCC
TGGCCGAGTTTTCACTCTGGCCTTTCGACTGATTCTCGGCTTCCCCCTGTACCTCTTTACCAACGT
GTCCGGAAAGAAGTACGAGCGATTCGCTAACCACTTTGACCCCATGTCCCCTATCTTCACCGAG
CGAGAGCGAGTGCAGGTCCTGCTCTCGGACCTGGGACTCATTGCCGTCGCTTACGCCGTTCGAA
AGGCTGTGCTGGCCAAGGGTGCCGCTTGGGTCATGTGTATCTACGGCGTTCCCGTGCTCGTTGT
GTCTGCCTTCTTTGTCCTGATTACCTACCTCCACCATACTCACCTGTCGCTCCCTCATTACGACTCT
ACCGAGTGGGACTGGATCAAGGGAGCTCTGTCCACTATTGACCGAGACTTCGGTTTTCTCAACC
GAGTCTTCCACGACGTTACCCACACTCATGTCCTGCACCATCTCATCTCGTACATTCCCCACTACC
ATGCTAAGGAGGCCCGAGACGCTATCAAGCCCGTGCTGGGCGAGCACTACAAGATTGACCGAA
CCCCTATCATTAAGGCCATGTGGCGAGAGGCTAAGGAGTGCATCTACATTGAGCCTGACCAGG
GCTCCGAGCACCAGGGAGTGTACTGGTACCATAAGCTGTAA

Translation of coding sequence for Epa955RACE-opt for *Yarrowia lipolytica:*

MGAGGRMTTASDDQKNILERVPCSKPPFSLADLKKAIPPHCFERSVIRSSYYVVHDLAVAYVFYYLA
NTYIPLLPSPFAYLAWPLYWFCQGSILTGVWVIGHECGHHAFSEYQWIDDTVGFVLHSALFTPYFSW
KYSHRNHHANTNSLDNDEVYIPKVKSKVTLYSKILNNPPGRVFTLAFRLILGFPLYLFTNVSGKKYERF
ANHFDPMSPIFTERERVQVLLSDLGLIAVAYAVRKAVLAKGAAWVMCIYGVPVLVVSAFFVLITYLH
HTHLSLPHYDSTEWDWIKGALSTIDRDFGFLNRVFHDVTHTHVLHHLISYIPHYHAKEARDAIKPVLG
EHYKIDRTPIIKAMWREAKECIYIEPDQGSEHQGVYWYHKL.

Nucleotide sequence for Epa955Trans-opt for *Y. lipolytica:*

ATGGGTGCCGGTGGCCGAATGACTACTGCCTCCGACGACCAGAAGAACATCCTGGAGCGAGTG
CCCTGCTCTAAGCCCCCCTTTTCCCTGGCCGACCTCAAGAAGGCTATCCCCCCTCACTGCTTCGA
GCGATCCGTCATTCGATCCTCTTACTACGTCGTGCACGACCTGGTTGTCGCCTACGTGTTCTACT
ACCTCGCTAACACTTACATCCCCCTGCTCCCCTCCCCTTTCGCTTACCTGGCCTGGCCTCTCTACTG
GTTTTGCCAGGGATCTATCCTGACCGGCGTCTGGGTCATCGGCCACGAGTGTGGTCACCATGCC
TTCTCTGAGTACCAGTGGATTGACGATACCGTCGGTTTTGTGCTGCACTCGGCCCTCTTCACTCC
CTACTTTTCGTGGAAGTACTCCCATCGAAACCACCATGCTAACACCAACTCCCTGGACAACGATG
AGGTGTACATCCCTAAGGTTAAGTCTAAGGTCACCTTCTACTCGAAGATTCTGAACAACCCCCCT
GGACGAGTGTTCACTCTCGCCTTTCGACTGATCCTCGGATTCCCCCTGTACCTCTTTACCAACGTT
TCTGGCAAGAAGTACGAGCGATTCGCTAACCACTTTGACCCCATGTCGCCTATCTTCACCGAGC
GAGAGCGAGTTCAGGTCCTGCTCTCCGATCTGGGCCTCATTGCCGTGGCTTACGCCGTTCGAAA
GGCTGTCCTGGCCAAGGGTGCCGCTTGGGTTATGTGCATCTACGGAGTGCCCGTTCTCGTGGTT
AACGCCTTCTTTGTCCTGATTACCTACCTCCACCATACTCACCTGTCCCTCCCTCATTACGACTCTA
CCGAGTGGGATTGGATCAAGGGCGCCCTGTCGACTATTGACCGAGATTTCGGTTTTCTCAACCG
AGTGTTCCACGACGTTACCCACACTCATGTGCTGCACCATCTCATCTCTTACATTCCCCACTACCA
TGCTAAGGAGGCCCGAGACGCTATCAAGCCCGTCCTGGGAGAGCACTACAAGATTGATCGAAC
CCCTATCATTAAGGCCATGTGGCGAGAGGCTAAGGAGTGTATCTACATTGAGCCCGATCAGGG
TTCCGAGCATCAGGGTGTCTACTGGTACCACAAGCTGTAA

Translation of coding sequence for Epa955Trans-opt for *Y. lipolytica:*

MGAGGRMTTASDDQKNILERVPCSKPPFSLADLKKAIPPHCFERSVIRSSYYVVHDLVVAYVFYYLAN
TYIPLLPSPFAYLAWPLYWFCQGSILTGVWVIGHECGHHAFSEYQWIDDTVGFVLHSALFTPYFSWK
YSHRNHHANTNSLDNDEVYIPKVKSKVTFYSKILNNPPGRVFTLAFRLILGFPLYLFTNVSGKKYERFA
NHFDPMSPIFTERERVQVLLSDLGLIAVAYAVRKAVLAKGAAWVMCIYGVPVLVVNAFFVLITYLHH
THLSLPHYDSTEWDWIKGALSTIDRDFGFLNRVFHDVTHTHVLHHLISYIPHYHAKEARDAIKPVLGE
HYKIDRTPIIKAMWREAKECIYIEPDQGSEHQGVYWYHKL.

Nucleotide sequence for Epa2161RACE:

ATGGGCGCAGGTGGTCGCATGTCGGACCCATCGGAGGGCAAAAATATTCTGGAACGAGTCCCG
GTTGGTCCACCGTTCACGCTAAGTGGTCTAAAGAAAGCTATCCCTGCACATTGCTTCGAGCGAT
CTGTCATCCGTTCATCCTACTATGTTGTTCATGATCTCATTGTAGCCTATGTCTTCTACTTCCTAGC
CAACACATATATTCCTATTCTCCCTACTCCTCTAGCTTACTTGGCATGGCCAGTTTACTGGTTTTG
TCAAGCCAGCATTCTCACAGGCCTATGGGTCATTGGTCATGAATGCGGTCACCATGCCTTTAGC
GACTACCAGTTGATCGATGACACTGTCGGCTTCATCCTCCATTCTGCTCTCCTCACTCCTTATTTC
TCTTGGAAATACAGCCACCGAAATCACCATGCCAACACAAATTCACTCGATAACGATGAAGTTT
ACATTCCTAAACGCAAGGCCAAAGTTGCCATTTACTCAAAGATGCTAAACAACCCACCAGGTAG
GGTGTTCACTTTGGTTTTCAGGTTGACTCTGGGATTTCCATTATACCTCTTAACAAACATTTCCGG
AAAGAAATACGGGAGGTTCGCAAACCACTTCGATCCACTAAGTCCCATTTTCACCGAGCGCGAG
CGGGTTCAGGTTCTTCTATCGGACCTTGGTCTTCTTGGGCTTTTCTATGCAATCAAGCTACTTGTA
GCAGCAAAGGGAGTTGCTTGGGTAATTTGCATGTATGGAGTTCCTGTGCTAGGCGTGCATATGT
TTTTCGTTTTGATCACGTATTTGCACCACACCCATCTCTCGTTACCTCATTACGATTCAACCGAGT
GGAACTGGATCAAAGGGGCCTTATCAACAATCGATAGGGACTTCGGATTCCTAAACAGGGTCC
TCCATGACGTTACACACACTCACGTCTTGCATCATTTGATCTCGTATATTCCACATTATCATGCAA
AGGAGGCAAGGGATGCAATCAAGCCAATCTTGGGTGAATTTTACAAGATTGATAGGACCCCAA
TATTCAAGGCAATGTGGAGAGAGGCCAAGGAATGCATGTACATTGAGCCGGATGAGGATAGC
GAACACAAAGGTGTCTATTGGTACCATAAAATGTAA

Translation of coding sequence for Epa2161RACE:

MGAGGRMSDPSEGKNILERVPVGPPFTLSGLKKAIPAHCFERSVIRSSYYVVHDLIVAYVFYFLANTYI
PILPTPLAYLAWPVYWFCQASILTGLWVIGHECGHHAFSDYQLIDDTVGFILHSALLTPYFSWKYSHR
NHHANTNSLDNDEVYIPKRKAKVAIYSKMLNNPPGRVFTLVFRLTLGFPLYLLTNISGKKYGRFANHF
DPLSPIFTERERVQVLLSDLGLLGLFYAIKLLVAAKGVAWVICMYGVPVLGVHMFFVLITYLHHTHLSL
PHYDSTEWNWIKGALSTIDRDFGFLNRVLHDVTHTHVLHHLISYIPHYHAKEARDAIKPILGEFYKIDR
TPIFKAMWREAKECMYIEPDEDSEHKGVYWYHKM.

Nucleotide sequence for Epa2161Trans:

ATGGGCGCAGGTGGTCGCATGTCGGACCCATCGGAGGGCAAAAATATTCTGGAACGAGTCCCG
GTTGATCCACCGTTCACGCTAAGTGATCTAAAGAAAGCTATCCCTGCACATTGCTTCGAGCGATC
TGTCATCCGTTCATCCTACTATGTTGTTCATGATCTCATTGTAGCCTATGTCTTCTACTTCCTAGCC
AACACATATATTCCTATTCTCCCTACTCCTCTAGCTTACTTGGCATGGCCAGTTTACTGGTTTTGT
CAAGCCAGCATTCTCACAGGCCTATGGGTCATTGGTCATGAATGCGGTCACCATGCCTTTAGCG
ACTACCAGTTGATCGATGACACTGTCGGCTTCATCCTCCATTCTGCTCTCCTCACTCCTTATTTCTC
TTGGAAATACAGCCACCGAAATCACCATGCCAACACAAATTCACTCGATAACGATGAAGTTTAC
ATTCCTAAACGCAAGGCCAAAGTTGCCATTTACTCAAAGATGCTAAACAACCCACCAGGTAGGG
TGTTCACTTTGGTTTTCAGGTTGACTCTGGGATTTCCATTATACCTCTTAACAAACATTTCCGGAA
AGAAATACGGGAGGTTCGCAAACCACTTCGATCCACTAAGTCCCATTTTCACCGAGCGCGAGCG
GGTTCAGGTTCTTCTATCGGACCTTGGTCTTCTTGGGCTTTTCTATGCAATCAAGCTACTTGTAGC
AGCAAAGGGAGTTGCTTGGGTAATTTGCATGTATGGAGTTCCTGTGCTAGGCGTGCATATGTTT
TTCGTTTTGATCACGTATTTGCACCACACCCATCTCTCGTTACCTCATTACGATTCAACCGAGTGG
AACTGGATCAAAGGGGCCTTATCAACAATCGATAGGGACTTCGGATTCCTAAACAGGGTCCTCC
ATGACGTTACACACACTCACGTCTTGCATCATTTGATCTCGTATATTCCACATTATCATGCAAAGG
AGGCAAGGGATGCAATCAAGCCAATCTTGGGTGAATTTTACAAGATTGATAGGACCCCAATATT
CAAGGCAATGTGGAGAGAGGCCAAGGAATGCATGTACATTGAGCCGGATGAGGATAGCGAAC
ACAAAGGTGTCTATTGGTACCATAAAATGTAA

Translation of coding sequence for Epa2161Trans :

MGAGGRMSDPSEGKNILERVPVGPPFTLSGLKKAIPAHCFERSVIRSSYYVVHDLIVAYVFYFLANTYI
PILPTPLAYLAWPVYWFCQASILTGLWVIGHECGHHAFSDYQLIDDTVGFILHSALLTPYFSWKYSHR
NHHANTNSLDNDEVYIPKRKAKVAIYSKMLNNPPGRVFTLVFRLTLGFPLYLLTNISGKKYGRFANHF
DPLSPIFTERERVQVLLSDLGLLGLFYAIKLLVAAKGVAWVICMYGVPVLGVHMFFVLITYLHHTHLSL
PHYDSTEWNWIKGALSTIDRDFGFLNRVLHDVTHTHVLHHLISYIPHYHAKEARDAIKPILGEFYKIDR
TPIFKAMWREAKECMYIEPDEDSEHKGVYWYHKM.

Nucleotide sequence for Epa2161Trans-opt for Y. lipolytica:

ATGGGAGCTGGTGGACGAATGTCTGACCCCTCTGAGGGAAAGAACATCCTGGAGCGAGTGCCT
GTGGACCCCCCTTTTACTCTGTCGGACCTCAAGAAGGCCATCCCCGCTCACTGCTTCGAGCGATC
CGTTATTCGATCCTCTTACTACGTCGTGCACGATCTGATCGTCGCTTACGTGTTCTACTTTCTGGC
CAACACCTACATCCCCATTCTCCCCACTCCTCTGGCCTACCTCGCTTGGCCTGTCTACTGGTTCTG
CCAGGCTTCTATCCTGACCGGCCTCTGGGTCATCGGACACGAGTGTGGCCACCATGCCTTCTCG
GACTACCAGCTGATCGACGATACCGTGGGTTTCATTCTCCACTCCGCTCTGCTCACTCCCTACTTT
TCGTGGAAGTACTCCCATCGAAACCACCATGCCAACACCAACTCTCTGGACAACGATGAGGTGT
ACATCCCTAAGCGAAAGGCCAAGGTTGCTATCTACTCGAAGATGCTGAACAACCCCCCTGGTCG
AGTGTTCACCCTCGTTTTTCGACTGACTCTCGGATTCCCCCTGTACCTGCTCACCAACATCTCCGG
CAAGAAGTACGGACGATTCGCCAACCACTTTGACCCCCTGTCCCCTATTTTCACCGAGCGAGAG
CGAGTTCAGGTCCTGCTCTCTGATCTGGGCCTGCTCGGTCTCTTTTACGCTATCAAGCTGCTCGT
CGCCGCTAAGGGAGTTGCCTGGGTCATTTGCATGTACGGCGTGCCCGTTCTCGGTGTTCACATG
TTCTTTGTCCTGATCACCTACCTCCACCATACTCACCTGTCTCTCCCTCATTACGACTCGACCGAGT
GGAACTGGATCAAGGGAGCCCTGTCCACTATTGACCGAGATTTCGGCTTTCTGAACCGAGTCCT
CCACGACGTGACCCACACTCATGTGCTGCACCATCTCATCTCTTACATTCCCCACTACCATGCCAA
GGAGGCTCGAGACGCCATCAAGCCCATTCTGGGAGAGTTCTACAAGATCGATCGAACCCCTATT
TTTAAGGCTATGTGGCGAGAGGCCAAGGAGTGTATGTACATTGAGCCCGACGAGGATTCTGAG
CATAAGGGAGTTTACTGGTACCACAAGATGTAA

Translation of coding sequence for Epa2161Trans-opt for Y. lipolytica:

MGAGGRMSDPSEGKNILERVPVDPPFTLSDLKKAIPAHCFERSVIRSSYYVVHDLIVAYVFYFLANTYI
PILPTPLAYLAWPVYWFCQASILTGLWVIGHECGHHAFSDYQLIDDTVGFILHSALLTPYFSWKYSHR
NHHANTNSLDNDEVYIPKRKAKVAIYSKMLNNPPGRVFTLVFRLTLGFPLYLLTNISGKKYGRFANHF
DPLSPIFTERERVQVLLSDLGLLGLFYAIKLLVAAKGVAWVICMYGVPVLGVHMFFVLITYLHHTHLSL
PHYDSTEWNWIKGALSTIDRDFGFLNRVLHDVTHTHVLHHLISYIPHYHAKEARDAIKPILGEFYKIDR
TPIFKAMWREAKECMYIEPDEDSEHKGVYWYHKM.

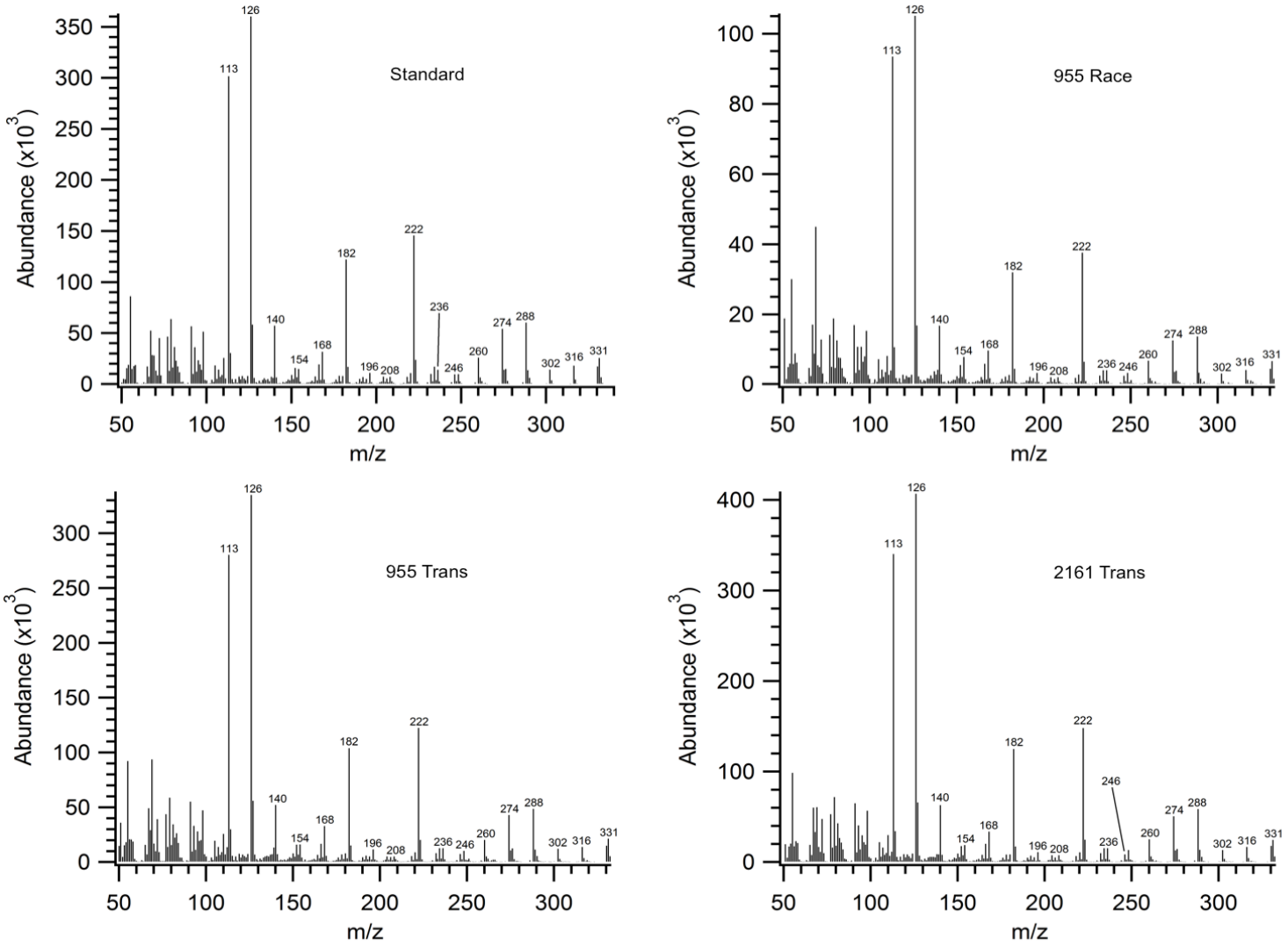


Figure 48: Mass spectral identification of crepenynoyl DMOX. DMOX mass spectra demonstrating the presence of crepenynate in yeast expressing Epa955 and Epa2161. Standard prepared from methyl crepenynate isolated from *C. alpina* seed oil is shown in the upper left panel.

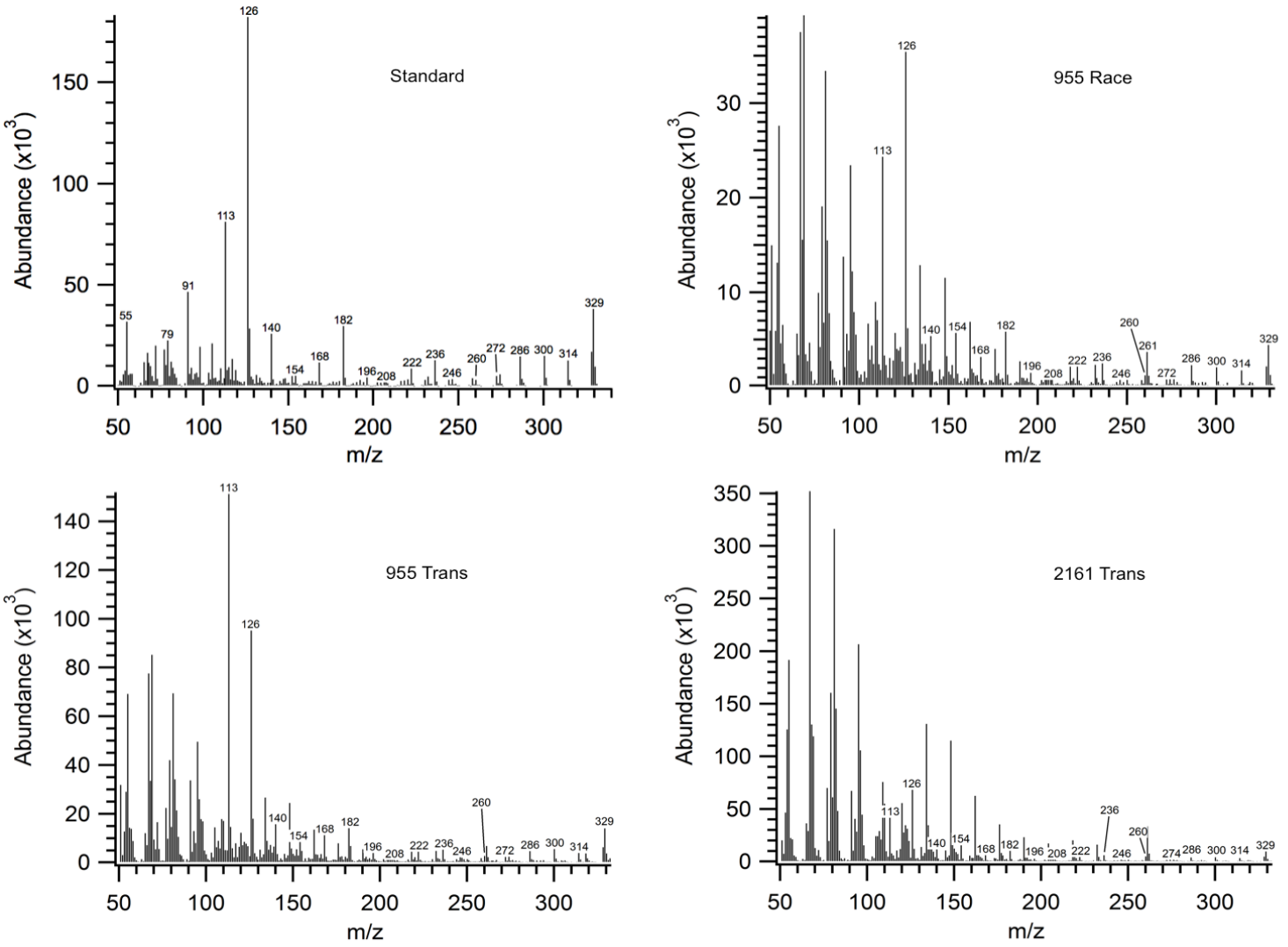


Figure 49: Mass spectral identification of dehydrocrepenynoyl DMOX. DMOX mass spectra demonstrating the presence of dehydrocrepenynate in yeast expressing Epa955 and Epa2161. Standard prepared from methyl dehydrocrepenynate isolated from *C. cibarius* is shown in the upper left panel.